

nature

UK's wasted opportunity at UNCSTD

Whatever may or may not be decided at UNCSTD in Vienna these two weeks, it is perfectly clear that the world scientific and technological community is experiencing strong pressures for change, for instance through proposals for the establishment of a massive global research and development fund, and through demands that the legal framework for the transfer of technology be restructured. It is thus vitally important that all those who attend UNCSTD take back to their home country a clear impression of the forces at work and alert their colleagues in every possible way to the need for new thinking. To a degree press reports help, but a very important ingredient is what individual people who attend do.

The constitution of the British delegation to UNCSTD is

— for whatever reason — such as to make this job very difficult to perform for the scientific and technological community. A delegation of twenty or so could have included some well connected scientists, some influential industrialists, some representatives of education, and some non-governmental development experts in addition to the necessary complement of civil servants. Instead the delegation comprises a minister (for two days), eight civil servants, five representatives of the Vienna embassy, three private or personal secretaries and only two people who could in any sense be regarded as outsiders. The burden on these two to inform the broadest community is intolerable. It will be interesting to see what help they get from their civil-servant colleagues in the months ahead.

Science, nonsense and responsibility

THE right — and responsibility — of the layman to exert an influence on the actions of scientists and technologists is becoming firmly established, as is evidenced this week by the controversy over the exclusion of journalists from an international meeting on short-term testing for carcinogens (see p.623). But is there not an equal right and responsibility for the scientist to guide the layman? Guidance is surely needed. The misuse of science by those often lacking in scientific training and understanding, but never lacking either conviction or commercial acumen, is now commonplace. It is commonplace because there are profits to be made from it.

For example, it is there even in the bookshops. A new form of escapism has been packaged as an alternative to science fiction, romantic fiction and thrillers. The new genre treats as fact a wide range of spurious topics. Themes include the artefacts left by our extragalactic forefathers, the possibility of disappearing through triangular holes in the sky over the Caribbean, the everyday lives of UFO-flying folk, paranormal fork-bending and outrageous claims about the influence that this additive-free food or that celestial body might have on our lives. A factor common to all these books, it seems, is the need for belief. A recent example of the extent to which strongly held but totally unfounded belief can resist the weight of fact is the saga of the alleged anti-cancer drug laetrile.

Can this need for belief be satisfied by something more sober than UFOs and cancer cures? By, say, a belief that an analytical and scientific approach to a physical problem will lead to as near as we can get to a solution? This will be possible only if genuine critical science, with its clarities

and uncertainties, is sold to the public with the determination that sells pseudo-science. It does not help to hold closed meetings and so increase the public's natural distrust of 'the expert'. The exercise is difficult, but it must be undertaken, or the voice of reason will not be heard at all.

True science does answer back sometimes. The BBC *Horizon* series carried an excellent critique of von Daniken's books. But many thousands more of his books will have been sold since then. Science should answer back more loudly and more often. One medium which invites such a response is the relatively new one (in the UK) of local radio, which thrives on the cheap-to-run 'phone-in' programme. In these programmes all forms of fringe and suspect forms of pseudo-science are often discussed by 'experts', and the more far-fetched the thesis proposed the more enthusiastic the response from the listeners. Just as harmful are 'open-line' programmes in which callers attempt to discuss subjects which neither the caller nor the programme presenter comprehends. Presenters are generally conscientious and, especially on medical matters, make pleas for calls from more qualified listeners. The pleas are often unanswered. This seems a pity because, although discussion of this type are certain to be shallow and are often not conducted very well, a shallow discussion with some sense in it must be better than one with no sense.

It may be true that 'truth will out', as it did eventually for laetrile. But scientific truth will 'out' much faster if it is sold with more enthusiasm, and if this means descending to the level of the market place, then perhaps that's as it should be.



The poor meet the rich

'Nature' correspondents report from Vienna on the opening sessions of the United Nations Conference on Science and Technology for Development



Third World: we will negotiate

DELEGATES to UNCSTD from a wide range of Third World countries agreed at a meeting in Bucharest last week that, at least at the start of the conference, they will bury economic and ideological differences behind a common negotiating strategy. They also agreed to stand firm behind the developing countries' proposals for a draft plan of action to emerge from UNCSTD — although they are prepared to negotiate on the plan's main demands rather than invite confrontation by sticking to a hard line.

The meeting was organised by the Group of 77 which has been negotiating on behalf of over 120 developing countries during the UNCSTD preparatory meetings. Representatives from the developed countries were relieved to hear that the Group of 77 would not seek initial confrontation. Although the developed countries are not prepared to accept all the demands that the developing countries have included in the proposed plan of action, they are keen, for political and economic reasons, to engineer a successful outcome to the conference.

Before the Bucharest meeting, some developing countries had complained that the proposed draft plan did not reflect their particular concerns. (The draft is the basis for the negotiations currently taking place in Vienna). Some African countries, for example, had complained that the proposals were too heavily weighted towards measures to benefit the more advanced of the Group of 77 countries — for example, with conditions governing technology transfer — and would not do enough to meet the technical needs of the poorest countries.

The Bucharest meeting agreed to recommend that the plenary session in Vienna establish a special working group to look at proposals for areas in which greater scientific and technical effort should be focused. In return the African groups agreed to support the Group of 77's proposals for a new funding system and new institutional arrangements for science and technology with the UN. Both proposals will be examined by UNCSTD's two main committees.

Members of the Group of 77 still differ,

however, over how the new institutional arrangements would operate to assure the politically weaker countries adequate control over research funds and their application.

A statement summarising the demands of the plan of action was issued at the end of the meeting. It says that inequalities in all fields of international economic relations are especially acute in science and technology. Despite the efforts of the developing nations there has been no real progress towards setting up a New International Economic Order with science and technology forming an integral part. The statement expresses alarm at the large fraction of global research expenditure devoted to defence and asks for "free and full access to scientific and technical knowledge, discoveries and innovations".

International cooperation should also play an important part in harnessing science and technology to the social objectives of the developing countries, the statement says. "For this, global and fundamental structural changes are required in the existing distribution of scientific and technological capacities in the world, in order to ensure increased participation by the developing countries in the quest for new scientific and technological knowledge".

In line with the suggested plan of action, the statement urges that international law governing technology transfer should be "restructured" in line with the goals of the developing countries who should be able to participate in financing mechanisms set up for scientific and technological development. The statement also urges support for measures to increase scientific and technical cooperation between developing countries. It says that practical measures "should be undertaken to support the cause of national liberation of peoples and territories under foreign occupation, control and domination, racism and racial discrimination, to enable them also to benefit from the application of science and technology for their development".

David Dickson

Arab group expel Egypt and establish Libya as spokesman

The decisions taken at the Tenth Conference of Islamic Foreign Ministers, held in Fez during April, are now being implemented by a strong 'Arab group' that has emerged at UNCSTD. The Islamic conference decided to boycott Egypt on all issues and support the Palestine Liberation Organisation.

In the past Egypt has assumed a natural leadership on Arab issues, in the field of science and technology, and this is the first time it has faced a determined and joint opposition to its efforts to lead the Arab group. Libya was unanimously elected to lead the group last week-end despite the fact that, nominated as a vice-president for North Africa, it was absent from its first meeting.

Alienation of Egypt from the Arab group has provided even stronger Arab support for the PLO delegation to UNCSTD. "We are determined to get a nation-state status for the PLO" a member of the Syrian delegation declared. At their first meeting, the group decided not to articulate an OPEC (or more appropriately, AOPEC) position. OPEC Arab countries will play a low-key role where we are allowed to do so" and will merge their identities within the Arab group.

The group also decided to give full support to the announced position of the Group of 77. Although the stance taken by the Group of 77 is a little too radical for some of the more conservative Arab states, like Saudi Arabia, Kuwait and the UAE, these states have been persuaded to support the Group of 77 as much as possible. Similarly Syria, which has shown considerable opposition to some declarations of the Group of 77, has decided to overlook its objections and support the decisions of the Arab group. Dr Shakir Fahm, Syrian Minister of Higher Education and leader of the Syrian delegation to UNCSTD says, "now is the time for the developing countries to show

their strength. This strength lies in unity”.

What the Arab group did not discuss is the financing of a development fund to help scientific and technological activities in the developing countries. Whether Saudi Arabia and other oil-rich Arab states will back such a fund to the tune of \$65 million a year for the first two years of the fund, as some suggested, they did not decide — though there is some indication that Saudi Arabia will give in to the US pressure. The Saudi aid to developing countries granted through the Saudi Development Fund, far exceeds this sum in an average year. The Saudi decision to support the fund, if it is forthcoming, will therefore be in line with their aid policies.

In addition to national efforts in such countries, international cooperation can and should play an important part, the statement says. “For this global and fundamental structural changes are required in the existing distribution of scientific and technological capacities in the world, in order to ensure increase participation by the developing countries in the quest for scientific and technological knowledge.”

Ziauddin Sardar

‘In everyone’s interest’ to transfer technology

The transfer of technology is not an issue of necessary conflict between developed and the developing countries, according to Crown Prince Hassan of Jordan, who shoulders the responsibility for social and economic planning in Jordan. Speaking during the general debate of the second plenary meeting at UNCSTD, Prince Hassan said that “despite short term disagreements and apparent disparities, in the long term, technology transfer is in everybody’s interest”. He warned, however, that technology “is not a commodity which can be copied to exaction. The transferred technology has to be appropriate to the needs of the recipient country”.

Many developing countries, said the Prince, lack the basic conditions for the establishment of agencies that promote and enhance technology. Furthermore, technology in the developed countries is based on a particular concept of entrepreneurship — that motivated by economic and commercial considerations. The Prince questioned the desirability of the developing countries adopting such a concept and argued for a “spirit of resourcefulness” or “appropriate entrepreneurship”.

The Prince pointed out that at present much aid is channeled towards the poorest members of the Third World, the decisions for granting aid being made on the basis of gross national product (GNP). But there is hardly any relationship, argued the prince, between GNP per capita and the levels of development.

Waldheim attacks ‘wasteful consumerism’

A major task of UNCSTD is to help generate universal acceptance of the need to direct the use of humanity’s scientific and technological potential to constructive ends, Dr Kurt Waldheim, Secretary General of the United Nations, told the opening session of the conference on Monday morning.

In a speech which reflected many of the ideas put forward by the conference secretariat during the preparatory process, Dr Kurt Waldheim singled out three areas in which he claimed much of human ingenuity and innovative ability had been misdirected into areas not beneficial for mankind: military technology; a type of economic growth which ignored damage to the environment or the social and cultural fabric; and the support of “wasteful consumerism in a world where famine and malnutrition are tragically present”.

Another motivating factor behind the conference was the enormous imbalance in research and development being undertaken internationally, he said, resulting in developing countries having to depend excessively on imported technology. This hampered the growth of indigenous skills more harmonious with local conditions and social or economic needs.

“This becomes a self-generating process and breeds over-all dependence. The starting point for breaking this circle is to enable the developing countries to gather and share the scientific knowledge so as to enhance their technological capabilities and accelerate their development”, he said.

Three elements were crucial in designing a successful programme of action: a conscious political will, increased funding for science and technology at the national and international levels, and the institution of an efficient mechanism for this purpose. Finally, directly echoing words used on many occasions by conference Secretary-General Mr Frank da Costa, Dr Waldheim said that the real divisions in the world were not between the north and south or east and west, but between those in favour of a passive continuation of the status quo, “which is prejudicial to all” and those in favour of dynamism, changes and innovation. “Let this conference signal the fact that science and technology can unite the developed and the developing countries in the common cause of the world’s future as a whole”, Dr Waldheim said. □



Talking of the brain drain, the Prince emphasised that the human dimensions of the problem, “which have a direct bearing on technological and scientific development in both labour-exporting and labour-importing countries”, are often overlooked.

Crown Prince Hassan suggested that the labour exporting countries should turn the situation to their advantage through policies that aim to utilise the experience of their nationals working abroad. “These should take the form of annual feedback programmes, where highly trained scientists and technologists working abroad would come back to their countries for varying lengths of time”.

Making an obvious reference to the proposed global development fund, the prince said that “we have to be careful before we decide to create yet another international agency providing yet another variation on the existing manner of solving an already identified problem”.

Ziauddin Sardar

Britain opposes creation of new mechanisms

A greater proportion of national and international aid budgets should be devoted to scientific and technological aspects of development, but there is no need at present for new financing mechanisms or institutional arrangements to do this. Such was the main message of the British delegation’s contribution to the UNCSTD plenary debate on Wednesday morning.

Mr Neil Marten, UK Minister for Overseas Development, told the assembled delegates that, if developing countries so wished, Britain would be prepared to devote an increasing share of its bilateral aid programme to scientific and technological activities in the Third World. He also announced that Britain is considering plans to set up a technology transfer centre to act as a general switch-

board for enquiries about British technology.

However on the two main issues which are likely to dominate negotiations during UNCSTD — the Group of 77's proposal for a new international financing system for science and technology, and for a new co-ordinating body within the UN — the British response was lukewarm at best, reflecting the position of other EEC countries.

Nor did the British show any keenness to discuss the heated topic of technology transfer or the conduct of multinational companies, and area where encouragement and support of existing channels and procedures is preferred to greater regulation and control.

"Experience shows the advisability of making progress through continuing change based on tested practices and procedures, rather than embarking on abrupt changes of direction. We have serious doubts about the value of any new financial mechanisms or the separation of funds for science and technology for development", Mr Marten said.

"Given that aid budgets are limited, the separation of funds is likely to drain off the resources available for development into different pockets. This would make it difficult for individual recipient governments to set their own priorities clearly and would lead to inflexible and inefficient programmes." □

Danes urge the need for compromise

In one of the most direct statements made on the opening afternoon of the conference, the head of the Danish delegation said that it was "imperative" that a compromise be reached on various proposals for financial resources to support greater scientific and technological efforts in developing countries, if the conference was to be made a success.

"My country is definitely prepared to carry its fair and proportionate share of the financing of science and technology" the Danish representative, Mme Lise Ostergaard told the conference, an indirect reference to proposals for a new fund largely financed by the developed countries.

● A further pledge for additional financial support came from the Swedish delegation. Pointing out that Sweden is already one of the few countries which contributes more than the 0.7% of its gross national product to international development previously agreed as a global target, the head of the Swedish delegation said that agreement had already been reached on a number of recommendations which, for their implementation, demanded increased financial resources.

Sweden was prepared "together with other countries and on the basis of a fair burden-sharing formula, to make a substantial increase of our contributions in support of international cooperation in the field of science and technology.

US supports 'reasonable ventures'

Speaking on the opening afternoon of UNCSTD, the chairman of the US delegation announced that the US was prepared to join "reasonable ventures" to strengthen worldwide scientific and technological cooperation.

This theme was echoed in a message from President Carter to the conference, in which he said that the US was willing to support "all practical endeavours" to help overcome problems such as food scarcity, the energy crisis, and the population explosion.

Although no direct comment was made in either pronouncement, the choice of words was partly a response to some of the proposals contained in the draft plan of action submitted to the conference by the Group of 77 — such as the setting up of a new international financing mechanism totalling \$2 billion by 1985 and \$2 billion by 1990 — which has previously been characterised by US negotiators as being "unrealistic".

US officials have not made it clear what they would support as a more "practical" alternative. However they indicate that the executive branch would be prepared to look favourably in the order of \$1.50 million over the next two years, although pointing out that no firm commitment to such a contribution can be made without the support of Congress.

In his speech to the conference plenary session, Father Theodore Hesburg, president of Notre Dame University and chairman of the US delegation, said that the global economy was not working as well as it should for either the poor or the rich countries, and that the patterns of worldwide technology generation, diffusion and utilisation lacked the cohesion that would incorporate and benefit the majority of people. "We have not yet found the right mix between scientific excellence and needed technologies" he said.

It was an imperfect global economic order that did not fully benefit from the

"robust and dynamic role" of international business and industry, and had not yet found "the right balance between the interests of private enterprise and of the developing countries" he said — an implicit recognition of the need to negotiate on issues such as the conditions for the international transfer of technology.

"The task of this conference is not one of restating the errors of the past but of weaving science and technology into the fabric of the future, the fabric of development. We need collaboration, not confrontation", Father Hessburgh said.

Among the specific proposals were that the US would take the initiative to bring together the operators of remote sensing satellites as systems. In addition, it was necessary for developing countries to have access to the leverage provided by scientific and technological information, if more just and equitable patterns of scientific and technological cooperation were to be achieved.

Referring to President Carter's description of science and technology for development as a "joint venture", Father Hessburgh said that the challenges faced by developed and developing countries alike made such a joint venture a global imperative. "The US notes, therefore, with pleasure the declaration of Bucharest (see page 620) in which the developing countries reaffirmed their willingness to work with a sense of urgency to assure the success of this conference", he said.

Science and technology should be used to help bring forth a threefold harvest: in the realm of reason, the realm of reality and the realm of rights. Efforts made in the real of reason were not a zero-sum game, in which the gains of those who seek equality and parity would automatically register as a loss for those who possess more. "In this realm we can prove the mutual benefits thesis — that advances in any part of the world are for the benefit of all", he said.

David Dickson

Canada pledges increase in joint projects

The Canadian government has agreed substantially to increase the budget of the International Development Research Center (IDRC) to allow it to support a joint projects between scientists in Canada and research groups in developing countries.

Addressing the UNCSTD plenary session on Tuesday, M. Martial Asselin, recently-appointed secretary of state for the Canadian International Development Agency (CIDA), said that although the precise amount of funds allocated to this purpose is yet to be decided, it would be in the order of the one-third of IDRC's current budget — ie about \$12 to \$14 million, an amount equivalent to 1% of Canada's total foreign aid.

Although under its initial terms of reference, IDRC is permitted to fund joint ventures of this nature — which from the Canadian side could involve university research groups or scientists from federal or provincial laboratories — the vast majority of the center's current funds are spent solely on research efforts in the developing countries.

The new commitment will, according to the Canadian Government, provide a direct way of linking up the needs of developing countries with Canada's own scientific and technological capacities. It has yet to be decided whether there will be an expansion Canada's foreign aid budget next year. □



Time to drown your sorrows?

US scientists find carcinogen traces in Scotch whisky

US scientists have found traces of nitrosamines, one of the most potent known carcinogens, in six brands of Scotch whisky. A few months ago, similar traces were found in various types of beer.

In both cases it appears that the nitrosamines are formed during the drying of sprouted barley malt. The levels detected are small, less than two parts per billion, according to Dr David H. Fine and Dr E. Ulku Goff, who carried out a study as part of a project funded by the National Science Foundation.

However levels as low as 10 parts per billion of nitrosamines are known to increase the incidence of tumours in laboratory animals. The beer industry has already said that, although little is known about the effects of such low doses on humans, it is taking steps to eliminate the presence of nitrosamines. It expects to have a solution "within a matter of months".

A spokesman for the Distilled Spirits Council of the US said last week that there was a marked discrepancy between the results of the NSF study and measurements reported last year by the World Health Organisation in Geneva, which had found nitrosamine levels of less than 0.2 parts per billion in Scotch whiskies. The council said this difference "may raise questions about the accuracy" of the new measurements.

But Dr Lawrence Garfinkel of the American Cancer Society doubted that such small amounts of nitrosamines would cause cancer. The most significant factor for him was that many drinkers were also smokers; a synergistic effect could account for the increased death rate from mouth and throat cancers that had shown up in studies of beer, wine and whisky drinkers.

The brands of Scotch whisky said by the NSF study to contain nitrosamines are: Chivas Regal, Black and White, J & B, Ballantine's Sandy Scot and Cutty Sark. Nitrosamines were not discovered in White Label whisky, nor in 46 other alcoholic beverages that were also studied.

Press banned from conference

The press is to be excluded from a meeting to evaluate short term, *in vitro* tests for carcinogenicity, following a ruling by the conference's organising committee. A decision not to invite journalists to attend the October, 1979 meeting in Atlanta, Georgia was taken when some of the participants voiced fears that the attendance of the press at the event would have an inhibitory effect and restrict the usual scientific discussions on such occasions. It has been decided, however, that in view of media interest in the subject, information on the conclusions of the conference will eventually be issued in December, at a press conference to be held in Washington.

Decisions to exclude the press from important scientific meetings are not without precedent. Recently, COGENE, the Committee on Genetic Experimentations of the International Council of Scientific Unions attempted to ban the reporting of the proceedings of a meeting on genetic engineering at Wye College in Kent. (See *Nature*, 278, 496, 1979). The directive announcing the ban was greeted with a storm of protest and COGENE were forced to reverse their decision. The criticism of a reporting ban could perhaps have been predicted, given the public interest in genetic engineering.

Carcinogens are also a subject which are much in the public eye, a fact given emphasis by the current proposals for their control being discussed by Government Regulatory Authorities in Europe and the United States. European regulatory authorities appear to be far more in favour of *in vitro* tests as a screening net for carcinogens than do their counterpart organisations in the US. The Environmental Protection Agency, for example, views these tests with some doubt.

Hazards at work 'kill 100,000 a year'

In a draft report to the President, the Toxic Substances Strategy Committee of the US Council on Environmental Quality claims that "more than 100,000 workers are believed to die each year as a result of physical and chemical hazards at work". The committee estimates that between 20% and 38% of all cancer may be caused, in part, by occupational exposure to carcinogens.

The report says prevention is the major key to controlling these environmental hazards and calls for the rigorous application of more than two dozen statutes for the regulation of toxic chemicals already on the books. It also advocated more legislation, including setting up a federal authority responsible for cleaning up hazardous waste sites.

More than 43,000 chemicals are listed by

In view of the importance of the short term tests, and concern about the variability in results reported from different laboratories testing the same chemicals, two scientists — Dr John Ashby of ICI's Central Toxicology Laboratory, Alderley Edge, Cheshire, UK, and Dr F.J. Serres, NIEHS, Research Triangle Park, North Carolina, U.S., decided to organise a carefully controlled study to evaluate short term tests. Ashby, with the resources and expertise of ICI to assist him, was responsible for the synthesis, coding and distribution of some 42 chemicals to be tested by 25 *in vitro* assays in various laboratories throughout the world (see *Nature*, 274, 740; 1978). The results of this valuable collaborative programme will be discussed at Atlanta in October.

The one notable omission in the list of participants in Atlanta is Dr Bruce Ames of the University of California. Ames, who developed the famous *in vitro* test for carcinogenicity which bears his name declined to take part in the Ashby/Serres study. The Ames test has, however, been evaluated in several laboratories along with other tests.

Few scientists involved with short term tests are unaware of the significance of the Atlanta meeting; Government Regulatory Authorities are similarly aware. Dr David Douglas, Head of the UK Health and Safety Executives unit for collating information on toxic substances referred to the Atlanta meeting as the one which would "take the lid off" the subject of short term tests.

The importance of the Atlanta meeting cannot be denied. It is this which makes the intended exclusion of journalists so controversial. It remains to be seen whether the organisers will have second thoughts and reverse their decision.

Alastair Hay

the Environmental Protection Agency as being subject to regulation. It may take years to determine whether or not a substance is hazardous and the complexities in reaching such a judgement are enormous. But the committee feels that the difficulties can be mitigated, to some extent, by greater cooperation between agencies involved in toxics-related research. The Government is to spend around \$1 billion on toxics research next year.

The report has drawn some criticism from the chemical industry. The Chemical Manufacturers Association called it a "hodgepodge of concepts and recommendations that have been either disputed or discredited since the committee was created in 1977."

Anne Norman

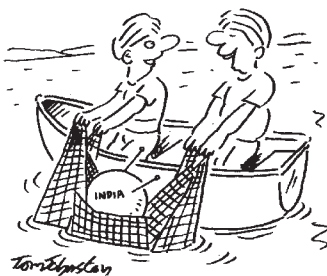
news in brief

Windscale issues emergency advice: Residents living within a mile of the UK Windscale nuclear fuel reprocessing plant in Cumbria have been issued with leaflets explaining emergency procedures in the event of any serious incident there. Similar booklets have also been sent to farmers within 10 miles of the plant. The move follows a recommendation by Justice Parker — whose inquiry report last year gave the go-ahead for expansion at the plant — that local residents be informed of all emergency procedures. At first explanatory leaflets were sent to local libraries and the media. Now these have been issued to the residents nearest to the plant to keep for reference. The booklets give details of evacuation plans, information broadcast procedures and medical arrangements in case of serious radiation leaks.

Harrisburg re-opening depends on residents' mental health: The mental health of residents living near the damaged Three Mile Island nuclear plant is a matter of "real and substantial concern" and should be carefully considered before allowing even the unaffected reactor on the site to be restarted, the US Nuclear Regulatory Commission has announced. Leonard Bickwit, NRC general counsel, said the commission felt that psychological, sociological and economic distress should all be considered in hearing the licensee's application to restart the plant, a decision unlikely to be taken by the NRC until the middle of next year. He added that a ruling on the mental health of residents could be a landmark in regulatory legal process.

This pronouncement contrasts with the recent publication by the operating company, General Public Utilities, of a plan to resume operations at the plant. This would involve venting into the air some of the radioactivity now contaminating the reactor chamber. However, although the quantities proposed would only involve releasing one tenth of a millirem of gamma radiation, compared with the 83 released at the height of the station emergency, it is unlikely the proposal will be given the go-ahead without strong legal and political opposition. And even if approved, it is unlikely that the plant could resume operations before June 1983.

Draft code to protect asbestos workers: A consultative document has been produced by the UK Health and Safety Commission on a proposed code of practice for workers dealing with asbestos insulation. The proposals are an extension and clarification of existing legislation. They include proposals for improving asbestos dust containment, waste disposal, personal protection and training for insulation workers. Industry and trade unions have until 30 November to give their view before a final code of practice is drawn up.



"Perhaps we'll have some benefit from our own technology now!"

Indian rocket crash: India's first rocket designed to put a satellite into orbit crashed into the Indian Ocean during a test flight. The launcher flew for only 15 minutes after blasting off from Sriharikota space research centre in South India. The test failure is a setback for the Indian space programme which has been preparing for greater independence from the

Soviet Union, whose launchers it uses at present. It was intended to put a satellite, Rohini Satellite-1, into orbit later this year using the SLV-3 rocket — which is similar to the US Scout rocket and uses solid propellants in all its four stages. Once the SLV-3 is successfully completed, India intends to launch a series of Rohini satellites — to be followed by a more ambitious programme of communications satellites.

Ariane delay: The first flight of Europe's independent space launcher, Ariane, has been postponed by about a month. The launch was scheduled for 3 November but because of problems with ball-bearing mounts for first and second stage engines, this has now been put back until the end of November or the beginning of December. However, a European Space Agency said the difficulty had not been a major one and had now been overcome.

African venereology conference opens: The first African International Conference on Sexually Transmitted Diseases was due to open at Ibadan University, Nigeria yesterday. More than 200 doctors and researchers were expected to attend and the opening speaker, Dr Duncan Catterall, described the event as "an historic day in African medicine". It had taken three years' campaigning to arrange the conference he said, and the agreement to stage it indicated that the African governments now recognised that sexually transmitted diseases were a major problem. Although no statistics were available, Dr Catterall believed the diseases were now widespread throughout the continent and there was an urgent need to develop a common policy towards treatment and contact tracing. In particular, the uncontrolled use of penicillin in many African countries was leading to the development of resistant strains which in future could prove very difficult and expensive to cure.

Two sponsors required for animal research under new Bill: Lord Halsbury's Bill on the protection of laboratory animals, due for its second reading in the House of Lords on 25 October (16 August, page 534), will require an applicant for a licence to do research on animals to have two sponsors, Lord Halsbury announced last week. One must be the applicant's supervisor, and the other an independent expert in the procedures to be used, who will also advise on alternative techniques. After the licence has been granted, the new licensee will have to work for a probationary period under the supervision of an experienced licensee. These rules will apply to those working with animals to produce vaccine and serums as well as those "experimenting" with animals.

The licensing of institutions where "procedures" on animals are carried out should ensure that stolen pets do not end up in laboratories, Lord Halsbury explained at a press conference on the Bill last week. The conference came to an abrupt end after interruptions from members of the Animal Welfare Society and the National Anti-vivisection Society.

Lord Halsbury's is not the only Bill on laboratory animals shortly to be considered by Parliament. Peter Fry MP will be putting another Bill to the House of Commons for its first reading shortly after the start of the next session.

US climate programme sets goals: Three major areas of activity — climate impact assessment, climate system research, and climate data, information and services — have been earmarked for the newly-established National Climate Programme, it was announced in the US. The programme was established by Congress last year and will coordinate the work of various federal, state, private, academic and international organisations. It is also planned that several experimental climate forecast groups should be set up as part of the programme, which will cost a total \$114.2 million next year. These will develop, evaluate and test new climate forecasting methods.

Soviet cosmonauts land: Cosmonauts Vladimir Lyakhov and Valerii Ryumin have returned to Earth after a record 175-day flight aboard the Salyut-6 space station. First reports say they are healthy owing to an "exceptionally conscientious" attitude to exercise — unlike the ground crew, who according to Flight Controller A.S. Eliseev "have been suffering a certain amount of weariness".

India: the struggle for useful science

Are India's scientists "spoilt brats" who need to be put in their place? Or are they being manipulated by power-hungry politicians? *Anil Agarwal* reviews the beleaguered state of science in India at the end of the two-year government of Mr Moraji Desai, who resigned as Prime Minister last month.

NEVER before was there more tension between the Indian government and the country's scientific establishment than during Prime Minister Morarji Desai's administration, which was catapulted into power in April, 1977, after the unexpected defeat of Mrs Gandhi at the polls. In October 1978, Mr Y.B. Chavan, the Opposition Leader in Parliament, took the unprecedented step of writing a protest letter to Mr Desai urging him to call an all-party meeting at the very earliest to discuss the problems created by the Janata Government's science policy. Mr Chavan said: "Science is a major national concern. It is not a party matter . . . our national commitment to science has been an agreed component of our national development policy right from independence". However, he alleged that since the Janata Party came to power, "several very disturbing developments have taken place . . . which have resulted in rapid demoralisation of the scientific community."

Mr Chavan criticised Mr Desai for permitting "foreign powers to make major inroads" into the country's nuclear programme and for unilaterally abjuring nuclear explosions even for peaceful purposes; for his totally unwarranted attack on the Council of Scientific and Industrial Research (CSIR) which in the face of a strong and unanimous reaction in the Parliament and the press, he had "to practically reverse" within barely three months; and, for mortgaging such vital sectors of the Indian economy as the aircraft and heavy electrical industries to Western and multinational interests.

Mr Chavan informed the Prime Minister that "we in the Congress Party believe that India's scientific and technological capability is a national resource, which any government whatever its political predilections, should cherish and develop to give strength to our country in a grimly competitive world."

For full thirty years, India was ruled by the Congress Party, first, under the leadership of Prime Minister Jawaharlal Nehru, and then, except for a brief period in the mid-sixties, under his daughter, Mrs Indira Gandhi. Both these leaders closely identified themselves with the scientific community. They made it a point to inaugurate every annual meeting of the Indian Science Congress; they would seldom miss an opportunity to visit a scientific laboratory or inaugurate a scientific seminar; and, they always kept

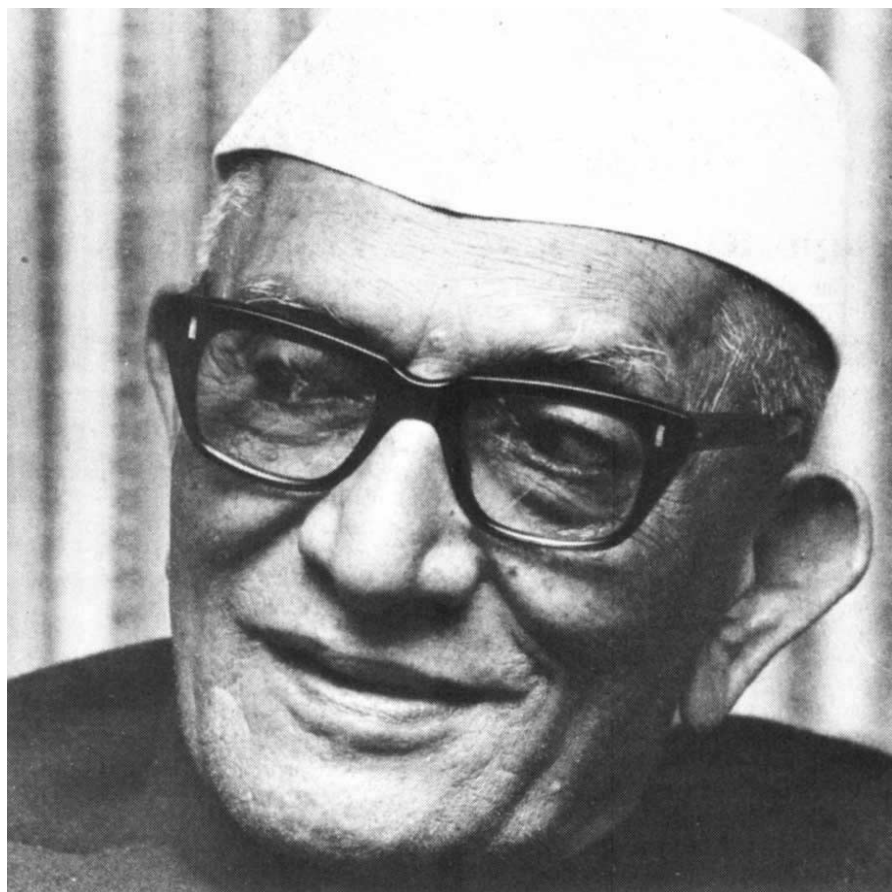
in their Prime-ministerial portfolio such scientific departments as atomic energy, space and electronics to ensure that the department did not suffer from any bureaucratic constraints. Mr Nehru also set the tradition that the President of the CSIR would be the Prime Minister.

He believed that science was the only tool that could help India to eradicate poverty without a violent revolution, and therefore, tried to give the scientific community enormous prestige. Mrs Gandhi would always describe Indian S&T capabilities as the bulwark of India's self-reliance. During their combined regimes, R&D funding increased 353-fold — from Rs. 11 million in 1948-49 to Rs. 3887 million in 1976-77 (1 Rs = £ at 1979 prices) — and giant organisations like the CSIR (with 44 scientific institutions), Defence Research and Development Organisation (with 31), Indian Council of Agricultural Research (29), Atomic Energy Commission (7), Indian Council of Medical Research (8), Department of Science and Technology (18), Space

Commission (5) and Electronics Commission (a funding and coordinating agency) grew up from scratch.

In 1977, the Central Ministries, public sector companies and State governments had another 327 R&D units; the private sector, 348; and 120 universities and institutes of technology were also conducting a variety of R&D projects. India had by then the world's third largest S&T manpower (only behind the US and USSR) consisting of 2.33 million, doctors, engineers, technicians and scientists. Over 53,700 professional were directly working on R&D projects in 1976-77.

Yet as far back as the mid-sixties, a feeling had started growing that India was not getting sufficient returns from this massive investment, particularly in industrial research. By the mid-seventies, Mrs Gandhi too occasionally began to chide Indian scientists for not conducting more relevant research. But the criticism of the Congress leadership was negligible compared to its praise. In a sense, it was a matter of vested interest for it to praise



Moraji Desai: some accused him of being anti-science

Indian R&D efforts, which, after all, as it rightly claimed, were there because of its own foresight. The vast urban intelligentsia, which grew up in the three decades of independence, (and which helped to sustain and drew sustenance from the Congress leadership) also came to see in these giant R&D establishments a sense of national achievement — a matter of pride for a developing country in a world where both political power and economic status essentially come out of the twin barrels of science and technology.

The new Janata Government was, however, not bound by this traditional association with Indian science, and, therefore, saw no reason to evaluate its achievements in muted tones. It immediately raised the question: have these organisations provided the country with benefits commensurate to the investments made in them? Dr Atma Ram, a former Director-General of CSIR, who was appointed by Prime Minister Desai as the chairman of the country's major science policy body, the National Committee on Science and Technology (NCST), explained in a recent lecture: "There is a vital difference between Morarji Desai's

period and the periods of his predecessors. During the period of his predecessors, most of the expenditure in S&T was devoted to building up the institutional infrastructure . . . This phase has now given place to the next phase of assessing benefits from science. People want to know how far they have benefited from the science in which they have invested their hard-earned money. It is natural that the Prime Minister should pose such questions before the scientists."

But why should there be any basis for asking such questions? In thirty years, Indian industry has grown substantially. India today manufactures computers, nuclear reactors, textile machinery, pharmaceutical plants, heavy engineering goods, tractors and other sophisticated industrial products. But the bulk of this manufacturing activity is based upon imported technology, obtained through foreign collaboration agreements which by 1977 totalled about 5200. Direct foreign investment has remained restricted as a government policy but know-how has been imported from Eastern Europe and the USSR as much as from Western multinationals. India's own R&D

institutions — even the national paper to the UN Conference on Science and Technology for Development admits — have played a very minor role in supplying the technology for this industrial development. The scientists and engineers in industrial establishments have themselves played a major and praiseworthy role in absorbing this imported expertise.

Take the case of Indian engineering goods for which the export trade has grown more than one hundred-fold — from Rs 52 million in 1956-57 to Rs 5517 million in 1976-77. These goods now constitute the most dynamic element of India's exports. India has also developed a series of its own mini-multinationals which have established several joint ventures with local enterprises in other developing countries. As the Indian government does not easily allow Indian companies to make foreign investments in cash, they have had to capitalise almost entirely on their know-how and equipment produced in India.

But why has the output of Indian R&D institutions been so poor? There are several reasons for this, and the main one is that Indian research institutions behave more like science laboratories than technology development institutions. They have failed to develop good links with prospective users of their know-how; they seldom carry their work beyond laboratory scale plants; most of them don't have adequate design and engineering capabilities; they seldom provide users with the software (raw materials procurement, credit, organisation, management, marketing, etc.) which is required to commercialise their hardware.

The Indian paper to UNCSTD admits in a milder language: "most Indian R&D agencies are much stronger in the generation of technology than in its delivery." But why is this so? Dr Atma Ram explains that "in those early days of the beginning of organised research efforts, there may not have been much understanding about the problems that would have to be dealt with in relation to the overall economic development of the country. There was perhaps a fascination for science, maybe a feeling of romanticism about it but not necessarily a clear and precise understanding of the mechanics of the application of science to economic growth and other spheres of national development. In other words, the necessary structure for research had been built with a noble but perhaps an undefined goal."

In March 1958, the Indian parliament had adopted a unique "Scientific Policy Resolution" (SPR) which was a declaration of intent by the people of India, as opposed to a resolution adopted by a particular government, to give sustained support to science. No other Third World country has ever made such a public commitment to science. The SPR was Prime Minister Nehru himself with some



T for Tata: assembly line in western India

Imported technology . . . exported trucks

One good example of an Indian mini-multinational is the Tata Engineering and Locomotive Company (TELCO), which is probably the first real Third World automobile multinational. TELCO even uses its own brand name. TELCO's joint venture in Malaysia assembles 7-ton 'TATA' trucks with components made by the parent TELCO in India. TELCO originally imported its technology from Daimler-Benz but now sells more than that company in Malaysia. It still does not have the capability to develop a totally new engine or a new vehicle. To manufacture a 13-ton truck in India, it has purchased a fresh licence from abroad. But through a well-organised technology assimilation programme, it extensively modified the 7-ton truck, concentrating on increasing its ruggedness and reliability — important attributes for bad road conditions in developing countries. Western companies meanwhile paid more

attention to speed, driver comfort and emission control, which are of greater concern only in the developed countries. TATA trucks are, therefore, in demand in the Third World.

Apart from its Malaysian joint venture, TELCO has also licensed assembly of its trucks to companies in Indonesia, Egypt, Guyana and United Arab Emirates. Assembly of scooters (of a design originally imported into India from Italy) has been licensed by another private sector company to Taiwan and Indonesia. Meanwhile, the premier public sector company, Bharat Heavy Electricals Ltd (BHEL), which in the past obtained its technology from USSR, Czechoslovakia, UK and West Germany, is now independently setting up power stations, boilers and transmission systems in Malaysia, Libya, New Zealand, Saudi Arabia and Jordan, and competing successfully with big companies abroad.



Should technology be backing heavy industry or traditional rural processes like sugar making (right)?

help from the Chairman of the Atomic Energy Commission, Homi Bhabha. But it is noteworthy that the word 'technology' does not appear anywhere in the text of the SPR, nor does it reflect the distinction between science and technology and the need for technological innovation.

It is very difficult to make a cost-benefit analysis of an R&D organisation like the CSIR which also conducts basic research and occasionally plays an important role in increasing the bargaining strength of Indian enterprises while purchasing foreign know-how. Out of the total number of processes developed in Indian R&D institutions, up to 1974-75, 1293 out of 1726 were from CSIR and out of the total number of 729 inventions, CSIR had 607 to its credit. (Incidentally, defence laboratories which remain outside the realm of public scrutiny, had only 46.) But many CSIR processes remain unused. The royalty earned from processes released for commercialisation does throw some light on the return made from investments in R&D. Total royalties earned by the National Research and Development Corporation, which offers technology developed in Indian R&D institutions for commercialisation, came to Rs 5.2 million in 1976-77. The CSIR alone spent Rs. 413 million in that year on R&D.

Entrepreneurs, both in the public and private sector, have generally claimed that they are afraid of investing in 'half-baked' CSIR technology. They, therefore, exert pressure on the government to allow import of know-how from established foreign sources. R&D institutions, on the other hand, have always complained that the technology they generate is not used by industrial establishments because of vested interests in importing from abroad. Often these entrepreneurs have even been allowed to re-import the same technology. This results in unnecessary remittances of scarce foreign exchange, breeds technological dependence, and reduces the demand for indigeneous technology. In this way, both the country and its scientific community suffer. In a bid to stop such repetitive imports and encourage the use of

local know-how, successive governments have allowed representatives of scientific establishments to participate in the various licensing committees, where they have the votes to stop any import of technology if it is already locally available, or it can be made available within reasonable time. But this practice has often led to delays in starting industrial establishments and is, therefore, not very appreciated by Indian entrepreneurs.

The Electronics Corporation of India Ltd, a public sector company that takes strong pride in indigeneous technology, was some years ago manufacturing and selling an outdated computer which was several times dearer than an equivalent computer abroad. The few components that had to be imported to manufacture it locally, were said to be higher than the total cost of a similar computer abroad. The Electronics Commission had to protest to the company about its price, as Indian users had no other alternative but to buy these computers because of a protected market. There is another case where a national laboratory offered to develop the technology for a public sector company to manufacture ferrites. The laboratory took several years to develop it, during which the country had to import the finished product in scarce foreign exchange. Finally, when the laboratory did supply the know-how, the company found that it still had to import the capital equipment.

Dr Atma Ram is a strong critic of unnecessary restrictions on import of know-how. It is "Made in India" and not "Invented in India" that matters ultimately for economic development. Indigeneous R&D, and import of technology, are not mutually exclusive, he says. "They are complementary. We need far more intensive efforts in R&D in areas in which technology is borrowed so as to adopt, modify and improve with the least possible delay."

Dr Atma Ram strongly criticised the restrictive approach to technology imports by the Electronics Commission arguing that the Indian electronics industry does

not even compare with that of far smaller countries like South Korea, Taiwan and Hong Kong. "We are not a technologically illiterate nation," says Dr Atma Ram. "Our experts can be depended upon to make the best available choice and build upon it, provided they are suitably utilised." Japan, USSR, South Korea and now even China, are all buying technology. "My fear is," he told me, "that it is not a militarised China but an industrially powerful China that will soon begin to threaten India." Professor M G K Menon, a close confidant of Mrs Gandhi, has since relinquished his position as Chairman of the Electronics Commission and is now Director-General of CSIR.

Desai's battle to reorganise research

But it is not the transfer of Professor Menon as much as the earlier attempt by Prime Minister Desai to reorganise the CSIR that has drawn the real ire towards the Janata Government. A few months after he had assumed office, Mr Desai suggested that the CSIR had become too large for one Director-General to manage. Without any semblance of consultations with the scientific community at large, he got the Indian cabinet to accept the reorganisation of the CSIR. Under this plan, 28 CSIR institutions were to be split off and attached directly to their main user ministries. Each R&D institution would have become a separate registered society with the relevant Minister as its President to ensure the autonomy of the institution. Some senior officials of the Ministry would also become members of its governing body. However, when the news leaked out, the reaction was hostile. The scientific establishment immediately saw in it an attempt to erode its autonomy and make it subordinate to interfering bureaucrats. Indian scientists have traditionally despised generalist bureaucrats for their power in decision-making. To circumvent bureaucrats and ensure direct access to Ministers, Mrs Gandhi's government started the practise of making

heads of most scientific organisations ex-officio permanent secretaries of their respective government departments. The Chairman of the Atomic Energy Commission is, thus, secretary to the Department of Atomic Energy as well. Mr Desai's move thus revived a traditional fear among scientists.

The scientific establishment further saw in the move to break up the CSIR, an attempt to sabotage the main source of the country's self-reliance. Several left intellectuals even described it as a move that would open up the country's industry to domination by Western multinationals. Some scientists argued that the new government was trying to "cut scientists down to size" because of the good relations they had enjoyed with Mrs Gandhi's government — which indeed was true even during the authoritarian emergency days, when the scientific establishment rushed to show its loyalty to its political benefactors, probably more than any other professional group. Others dismissed the new government as basically anti-science, and conveniently used the government's interest in simple rural technologies and Mr Desai's own interest in traditional medical systems like auto-urine therapy, to illustrate their point.

It was soon clear that Mr Desai's government had not only ineptly handled the reorganisation proposal but also totally misjudged the symbolic national value that the CSIR held in the minds of the urban intelligentsia. It was a rare newspaper that supported the government. Several members of parliament also protested. Headlines screeched: "CSIR has been an eyesore to the multinationals" and "CSIR is the first martyr to the Janata Government's ill-concealed solicitude for the international giants."

As is usual in Indian politics, such charges are seldom accompanied by proof. While it is possible that the proposal could result in more harm than good — through greater bureaucratisation of the laboratories linked directly to Ministries and thus greater stifling of their scientists — it is difficult to imagine that Prime Minister Desai could have deliberately taken a decision to sabotage the country's self-reliance. I interviewed a series of senior scientists, many strongly opposed to the government's proposal, but none would say that Desai had wrong intentions. All agreed that the output of organisations like the CSIR needs to be both increased and improved.

However, the proposal was strongly supported by the bureaucracy because they could have surreptitiously helped those Indian interests who prefer foreign technologies and thus increase the influence of Western multinationals, explained a senior scientist working in one of the government departments. Once scientists came under the direct control of bureaucrats, they could no longer oppose their wishes. And it is among the

bureaucracy, he alleged, that foreign technology-oriented industrial interests have considerable influence. The case of the research centre, that has traditionally been under the Minister of Posts and Telegraphs is cited. This centre has apparently done less to generate indigeneous technology and more to "sanctify the import of foreign technologies".

In the face of the unexpected political clout that the scientific community was able to muster, the proposal had to be scaled down. Finally, only four research institutes and, ten research associations were split off from the CSIR. The three museums under it were also transferred to the Ministry of Education. And it was decided that the reorganisation would be treated as an experimental measure, the performance of the separate institutions watched, and if so required, the decision could be reversed.

Getting more out of India's labs

But whatever may have been the outcome of Mr Desai's proposal, it is important to note that his original question remains: how can Indian laboratories be coaxed to produce more? His was not the first proposals for reorganisation. Several committees like the Administrative Reforms Commission and Sarkar Committee of Enquiry into CSIR have also considered reorganisation. But none of the proposals have ever met with success. The scientific establishment itself has produced no proposals, nor even made any effort to come up with some.

The Indian scientific organisations are themselves extremely hierarchial and bureaucratic establishments, where success is often the result of patronage. Within these organisations, it is hard to find any semblance of enquiry or debate. Whatever may happen, the working scientists remain quiet. In fact, there is no evidence to show clearly how the scientific community (as distinct from the top establishment) reacted to Mr Desai's proposal. "Must scientists keep mum?", asked Surendr Jha, editor of *Science Today*, India's leading science journal, in an editorial. The government today provides nearly 90 per cent of Indian R&D funds. These funds have brought with them a financial and administrative culture, that has resulted in a stifling bureaucratisation of the R&D institutions. Scientists have turned into scientist-bureaucrats, appendages of a vast system of patronage and hierarchy. It is as if these big laboratories are like the Taj Mahals of Indian science, built to glorify and bury it.

In this bureaucratic system, the director-generals of the various councils and chairmen of the various commissions enjoy vast power and they take decisions without any reference to their own area of specialisation. Ward Morehouse, a US

social scientist who has extensively studied Indian science policy, claims that these seven or eight scientists take virtually all the decisions in the world of Indian science. Committees are rampant in the decision-making procedure, and they are freely used to legitimise any whim or fancy by simply ensuring proper selection of their members — a practice that the Indian administration, in general, has imbibed in more than a full measure from Britain. Taking all the members of the NCST, directors of various laboratories, etc., decision-making of any national significance in science-related matters is restricted to about 200 individuals. The decisions made by the rest of the scientific community, if they make any, are of, at best, very local significance only.

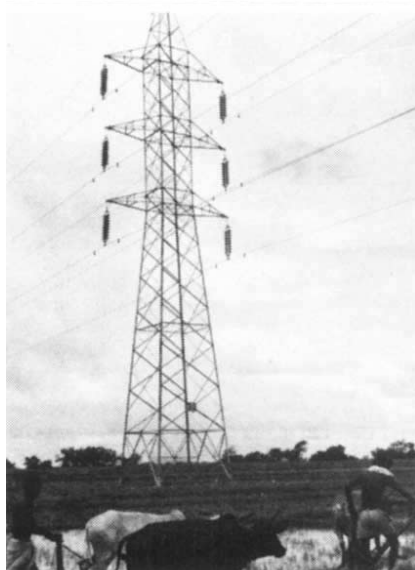
It is not surprising that in this system, power and prestige usually comes to an Indian scientist when he leaves his laboratory, stops any R&D work directly, and turns into a science administrator. A study of Indian scientific and technical manpower engaged in R&D recently released by the Institute of Applied Manpower Research, New Delhi, reveals that only nine scientists out of 23,700 surveyed had a salary above Rs2,500 and 108 were in the scale between Rs2,250-2,500. Almost all of them were administrators. The apex for personnel engaged in full-time research lay below Rs2,500.

It is, therefore, easy to appreciate that such powerful and hierarchical establishments will strongly resist any reorganisation, whittling down of their empires, or even any effective science planning, for that matter. The best example of resistance to the latter is provided by the work of the first National Committee on Science and Technology, which was constituted in 1971 as a small and compact body of 10 members chaired by Mr C Subramaniam, then Minister of Planning and Science and Technology. The principal task of the NCST was, in the words of Mr Ashok Parthasarathi, then Mrs Gandhi's special assistant on science and technology, "to prepare a comprehensive science and technology plan, derived from and closely integrated with the fifth Development Plan (1974-79). This was the first time that a deliberate political commitment had been made to integrate the economic, scientific and technological and educational dimensions of the development strategy into a coherent whole."

The members of this NCST were all young, middle-rank scientists, deliberately chosen to bring a breath of fresh air into science planning. But the giant organisations controlling agricultural research, defense research, atomic energy and space research, which consume the bulk of institutional R&D funds, made no serious attempt to co-operate. None of them wanted their areas of competence to be subjected to scrutiny by any committee outside. The NCST, too, recognising that it

Power plant deal that sparked off protests

One outstanding incident that attracted considerable public attention during Mr Desai's administration was the deal of the Bharat Heavy Electricals Ltd (BHEL) with the West German multinational, Siemens, to acquire know-how for a comprehensive range of heavy electrical equipment, including 200-500 MW coal-based power plants. The deal was expected to last 15 years, an unusually long period as the government generally frowns upon long-term collaboration agreements. Over 100 MPs and virtually every newspaper columnist again protested that this deal would make the country's electrical industry totally subservient to the foreign company. Part of the criticism emanated from pro-USSR MPs, who are piqued that BHEL was dumping technology that was originally acquired at easy terms from USSR. But many Indian public sector companies, who initially obtained know-how from socialist countries (often at critical moments when they had been totally denied access to know-how by Western sources) have in recent years sought collaboration with Western suppliers. This is because their technology demands have become more sophisticated, they themselves have



gained more bargaining strength, and because they are now looking for competitive export markets.

The Minister of Industries, George Fernandes, replied to the criticisms by pointing out that it was the Janata Government which sent IBM and Coca-Cola out of the country. His Ministry had promoted a drink developed by the Central Food Technology Research Institute, a CSIR constituent, to replace Coca-Cola. Moreover, the negotiations for the BHEL-Siemens deal had begun when Janata Ministers were behind bars, and therefore his government should not be described as "selling the country out to

multinationals". He himself supported the deal as one that would be useful to BHEL. The BHEL board approved it and so did the Ministry of Industry and the inter-ministerial Foreign Investment Board. But when public pressure mounted up, he had to seek the approval of the Indian Cabinet. And when the Cabinet couldn't reach a decision, it was referred to a specially constituted committee of five Cabinet Ministers for further consideration. It is not certain now whether the deal will go through.

The BHEL-Siemens deal can certainly be criticised for its comprehensiveness as it does appear to leave little to challenge BHEL's own engineers. It could have been more selective and BHEL's know-how demands could have been tempered by the company first seeking to fully supply the Indian market instead of being excessively worried about gaining more export muscle. Many vital industrial areas like West Bengal and Maharashtra are facing crippling power shortages. But the manner in which the deal has been questioned publicly gives the impression that many people believe Indian companies should be totally self-reliant for their technological needs. Except USA, there is no country in the world which generates itself even 10 per cent of the technology it uses. It is ironical that the atomic energy programme, which is held up as India's finest example of self-reliance, is really an example of a programme that is based on assiduously assimilated foreign know-how.

did not possess the muscle to take them on, concentrated mainly on the areas of R&D for which the CSIR is responsible. This too became possible apparently because the chairman of NCST was also then Minister in-charge of CSIR. As soon as the plan was prepared, the NCST was reconstituted and the young scientists replaced by the heads of the giant organisations. The simple explanation given was that the new NCST should rightly consist of the organisational heads as they are the people responsible for implementing the plan. But was that really the case? As the paper to UNCSTD puts it, the plan soon became "a useful check-list of programmes and projects", and that "suitable mechanisms for their implementation, monitoring and review could not be devised till the middle of the fifth plan (1974-79)." Some 2000 scientists had been involved in the preparation of the NCST plan, which was so easily, more or less, scuttled.

The next chairman of the NCST, Mr P N Haksar, seldom convened the reconstituted NCST, and, in fact, after a long delay, when he did, he asked it to discuss for a whole day whether there was a need for such a committee. As he put it, there were already so many barons to take care of Indian science. The NCST was, thus, soon reduced to a dormant state.

Until, of course, Mr Desai appointed Dr Atma Ram, a close confidant, as the third chairman, and then unpleasant decisions like the reorganisation of the NCST and review of the Electronics Commission began to be taken.

Prime Minister Desai certainly refused to acknowledge the need for the all-party meeting suggested by Mr Chavan and challenged "anyone to prove that the transferred institutions are not closely and intimately connected with the policies and programmes of the Department concerned." "I would go so far as to say," he pointed out, "that the scheme of reorganisation promotes the country's commitment to using science for national development." But despite this bold face, the Janata Government became increasingly defensive. Both Prime Minister Desai and Dr Atma Ram, chairman, NCST repeatedly pointed out that the Janata Government was neither anti-science or anti-scientists, and that actually increased the R&D budget from Rs3,887 million in 1976-77 to an estimated Rs4,767 million in 1978-79. The Prime Minister's speech to the Indian Science Congress in 1979 is reported to have been largely drafted by Professor Menon instead of Dr Atma Ram.

The lack of communications between the government and the Indian R&D

community has come at an inopportune time. Something that Indian scientists have always wanted, an R&D culture in the industrial establishments themselves is now steadily growing. Many large public and private sector companies have set up their own R&D units. Science organisations like the CSIR now not only need to increase and improve their output, but also begin to co-ordinate their work with these units and even slowly begin to reconsider their long-term role in Indian R&D. As industries begin to undertake routine industrial R&D themselves, R&D institutions could possibly move into more sophisticated and forward-looking basic and applied research that is required to support industrial R&D. In the present state of disarray, there appear to be very few science planners thinking at all, particularly in the scientific establishments.

The incidents during the Desai administration show that the scientific establishment can muster enormous political clout in a developing country. But, unfortunately, in the Indian case, it is behaving more like a pampered and spoilt super-brat, who immediately turns nasty when its ears are tweaked. But questions of public accountability in India will certainly arise again. □

correspondence

Fast reactor safety

SIR, — In his article on fast reactor safety (26 July, page 270) Norman Dombey claims to introduce to non-specialists some features of fast reactors that are not available outside the technical literature. The non-specialist would do well to treat this article with caution as it contains a number of errors (some of them elementary) and is more likely to mislead than to inform. The following is a list of some of the more outstanding errors:

- "Thermal reactors are designed to be in their most reactive nuclear configuration". This is incorrect and some examples were given by Professor F.R. Farmer (12 April, page 593). Some reactivity coefficients in thermal reactors are positive and are not all negative as stated — indeed gas-cooled Magnox reactors have operated safely for over 20 years with strongly positive moderator temperature coefficients.

- "Thermal reactor has less than one critical mass". This is incorrect: all reactors have to have at least one critical mass.

- "So in a fast reactor of the design envisaged in CFR1 . . . there will be no overall negative Doppler coefficient". The reasoning and the conclusions are both incorrect and there is in fact a strong negative Doppler coefficient which is a major contributor to the safety of the reactor. (I have now seen the erratum (9 August, page 444) which corrects this point.)

- "For example a coolant channel gets blocked, sodium boils . . . pieces of plutonium fall and form a molten mass which if large enough can become prompt critical". This is incorrect. The reactivity effect of voiding a single channel is so small that there would be little effect on the power level in the rest of the core. There is insufficient fuel in a single fuel sub-assembly to produce criticality even if all the fuel pins in it were melted and compacted.

There are other incorrect technical statements: for instance, sodium is more easily contained than high pressure water or steam and is not as corrosive to reactor constructional materials; and the delayed neutron fraction, on which the normal control behaviour of thermal reactors as well as fast reactors depends, is nearly twice as great as stated (and is always greater in a fast than in a thermal reactor with the same fissile material).

These errors are important in that they convey the impression that fast reactors are difficult to control, unsafe and not well understood. On the contrary the fast reactor in normal operation has no tendency to become unstable and can be left for hours without any adjustment of the controls. Temperature and power control is simple, straightforward and easy to engineer to very safe standards. Their safety has been the subject of intense study for very many years and it is well understood. Clearly there will continue to be work related to the detailed choice of safety features until licensing rules have been formulated and as with other reactor types the preferred selection of intrinsic and engineered safety features will be clarified during negotiation of the licence for the first commercial reactor. What is important is that fast reactor designers are confident that they can meet safety standards at least as stringent as those applied to other reactor types. Furthermore any design features incorporated for safety reasons are not likely to increase the expected cost by more than a few per cent.

The economic statements are equally ill-founded. The increases in the capital cost of

SNR 300 were due largely to late changes in the design requested during the complex licensing process which was being developed using SNR 300 as a guinea pig; even if the changes incorporated in this way were thought to be necessary for later reactors the cost of including them *ab initio* would be small compared with the costs of SNR 300.

Whilst the capital cost for a fast reactor appears to be inherently higher than for a thermal reactor — estimates range from as low as 1.1 times to about 1.5 times or perhaps approaching 2 times for the first-off commercial sized demonstration plants — there is little doubt that the lower fuel cycle costs will more than compensate as uranium prices increase, due to pressures in the next few decades on the limited supplies of uranium ore. The breakeven point could well be before the end of the century, which is the timescale for planning exploitation of the fast reactor system. Remember the fast reactor will produce about 60 times more energy than a thermal reactor for the same quantity of uranium. Commercial designs are of course being developed to meet safety requirements which are included in cost estimates, and the implications by Dr Dombey that designers have to suddenly take them into account after the rest of the design is complete shows a lack of understanding of the attitude of design engineers to safety.

Despite his failure to reach it by logical argument, Dr Dombey's general conclusion that the UK should continue to collaborate as closely as possible with other countries in developing the fast reactor is commendable and has been stated on many occasions to be current policy. Articles which undervalue the extent of the UK's possible contribution to such a collaboration are not likely to be of assistance in such collaborations.

It is disturbing that the article contains so many inaccuracies, which, compounded with the error about the Doppler Coefficient, gave such an unbalanced view of fast reactor safety at a time when an impartial academic view would have been particularly valuable.

Yours faithfully,

R.D. SMITH

Fast Reactor Development Directorate,
UKAEA, Risley, UK.

Domestic animals are not inferior

SIR, — Mellanby (*Nature* 275, 82; 1978) attacks opponents of 'factory farming' for their errors in animal behaviour, as well as for their own behaviour at his talk. First, while not wishing to justify the human behaviour he describes, it should be borne in mind that many in the environment movements have suffered far worse behaviour at the hands of establishment figures — treatment not all that different from what is meted out to dissidents in the socialist world, well detailed in the columns of *Nature*.

Second, there are scientific, practical and humane reasons for objecting to some aspects of intensive husbandry. Those who are critical of such practices do not get their views published — in contrast to the regular columns in *Nature* by such strong advocates of 'factory farming' as T. H. Jukes and Mellanby.

Third, since Mellanby frequently (and rightly) complains of occasional errors and oversimplifications made by environmentalists may we have the opportunity of correcting some of his.

The statement by Mellanby, "Domestic animals, cattle and sheep, and pets such as dogs and cats only exist because man breeds them . . . but if man disappeared from the world, so would most domestic animals" seems at variance with his later comment about the feral cat "... endangering the native fauna ...". Both statements are misleading.

The feral cat in Australia largely has the rabbit as its major prey (Coman, B. J. and Brunner, H., *J. Wildlife Management* 36, 848; 1972). Habitat destruction by man is the major threat to Australian native fauna (Marshall, A. J., *The Great Extinction*; Hienemann, Melbourne 1966).

In reviewing literature on the success and failure of introduced animals, we find that there is no statistically significant difference between domestic and wild species in regard to successful establishment in the absence of human care. Indeed, some major animal domesticates, notably goats, pigs, cats, dogs, and rabbits become successful in a variety of habitats and resist sustained human efforts at control or eradication.

The idea that domestic animals are inferior and degenerate — and thus less deserving of humane attention — is widespread. It is dangerous to confuse science and value judgements, however, and it is worth bearing in mind that Konrad Lorenz justified his support for the German 'eugenics programme' on the basis that man was becoming domesticated (Haldane, J.B.S. in *Culture and the Evolution of Man*; Galaxy Books, Oxford University Press, New York, 1962 and Nisbett, A., *Konrad Lorenz*; J.M. Dent, London 1976).

Yours faithfully,

C. M. ANN BAKER
CLYDE MANWELL

Selby, Yorkshire, UK.

Brazilian official was illegally dismissed

SIR — I wish to congratulate you for the story on the Brazilian exiled scientists (26 July page 268). I would only like to rectify a misquotation concerning my husband, an ex-official of the Brazilian Parliament who was illegally dismissed for political reasons when the military took over in 1964, and was not "an ex-deputy whose mandate was annulled", as published.

Yours faithfully,

ALICE RIVERA

CNRS, Gyf sur Yvette,
France.

What use is scientific awareness?

SIR — your concern for the lack of scientific awareness of the British public is commendable (9 August, page 435), but may I ask to what end? As an unemployed maths teacher who has never held any scientific or technical position, I am probably as close to the "general public" as your regular readership extends. I try to keep myself informed on all branches of science, and my general scientific awareness is far greater than could ever be reasonably expected as a norm. I think it is worth noting that as far as I know this awareness has never been of the slightest benefit to anyone.

Yours faithfully,

R. J. SALISBURY

Anglesey, Wales, UK.

news and views

Primitive model for cell cycle control

from Clive Lloyd

IN non-motile cells of multicellular organisms atypical, internal cilia form from the centriole. The role of these so-called 'primary' cilia in such cells is unclear but recent work shows that nevertheless they may provide a useful indicator of the state of the cell cycle.

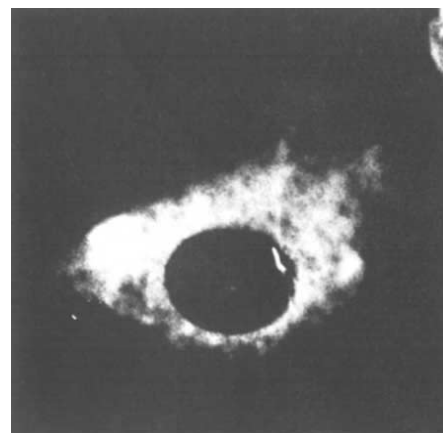
The centrioles appear in different guises at different stages in the eukaryotic cell cycle. During the division of most animal cells for instance, they are associated with each pole of the mitotic apparatus. They may, however, merely be passengers during this process, since amorphous pericentriolar material seems to be the major nucleation site for those microtubules formed at the poles of the mitotic apparatus (Gould & Borisy *J. Cell Biol.* **73**, 601; 1977). By contrast, the role of centrioles in non-dividing cells is less ambiguous for they act there as pre-formed templates (basal bodies) for the outgrowth of the microtubules which form cilia and flagella. Cilia and flagella are generally thought of as organelles of motility; propelling unicellular organisms through the environment or moving the environment over the surfaces of tissues (see *Cell Motility* (eds Stebbings & Hyams) Longmans, 1979). However, atypical, internal cilia are also seen associated with the centrioles in non-motile cells of multicellular organisms.

These internal cilia differ from the motile variety in several ways. Motile cilia/flagella generally consist of a bundle of nine outer doublet microtubules surrounding two singlet microtubules, but the internal cilia have a 9+0 instead of a 9+2 arrangement (Barnes *J. ultrastruct. Res.* **5**, 453; 1961). In addition, the 9+2 cilia possess a basal organisation involving only one centriole or basal body, whereas cilia possessing twinned basal structures generally display a 9+0 pattern (although the converse may not necessarily be true). The 9+0 cilia have been referred to as 'primary' since Sorokin (*J. Cell Sci.* **3**, 207; 1968) first noted that they were produced in bronchial epithelial cells before the 9+2

type which make up the ciliated border.

As for their function, primary cilia have been observed at the site of regrowth of cytoplasmic microtubules in 3T3 fibroblasts (Osborn & Weber *Proc. natn. Acad. Sci. U.S.A.* **73**, 867; 1976) (see figure) and are aligned with the current track of migration in such cells (Albrecht-Buehler *Cell* **12**, 333; 1977). This could imply some function for the primary cilium in influencing cell polarity, but Albrecht-Buehler is careful to point out the difficulty of sorting out cause from effect in such observations. Perhaps primary cilia serve no useful purpose. Found in unicellular and multicellular organisms alike, the centriole may be part of a primitive motile system and in differentiated, non-motile tissue cells, the 9+0 cilium on the centriole may be vestigial — representing a throw-back to a unicellular, free-swimming mode of existence.

However, even if it is now redundant, the cyclic formation and disappearance of the primary cilium still provides clues to cell cycle control. As long ago as 1898, it was suggested that the transformation of centrioles into ciliated or flagellum-bearing basal bodies could be one of the control mechanisms for the regulation of mitosis and differentiation (Heneguy *Arch. Anat. Microscop. Morphol. Exp.* **1**, 481; Von Lenhossek *Verh. Anat. Ges. Kiel.* **12**, 106). This theme was revived by Rash *et al.* (*J. ultrastruct. Res.* **29**, 470; 1966) who, studying cardiac differentiation, reasoned that "the blockage of mitosis usually associated with the initiation of cytodifferentiation can be correlated with the formation of cilia and may be mediated by the transformation of the mitotic centrioles into ciliary basal bodies." Several reports agree with this: primary cilia do not occur in mitotic BHK fibroblasts, but reappear soon after division (Archer & Wheatley *J. Anat.* **109**, 277; 1971); they are more frequently observed in differentiated, non-dividing brain tissue than in dividing tumours or juvenile tissue (Dingemans *J. Cell Biol.* **43**, 361; 1969); primary cilia are observed in rat liver cell cultures only after their growth has been arrested in the post-mitotic phase



Immunofluorescence micrograph of 3T3 cell stained with antibodies against tubulin. The cell has been treated with colchicine to destroy the microtubule network and the primary cilium is clearly visible. Magnification $\times 900$. (From Osborn & Weber *op. cit.*)

of G_1 , and not before (Mori *et al. Expl. Cell Res.* **120**, 435; 1979).

Now, in a recent paper (*Cell* **17**, 527; 1979) Tucker, Pardee and Fujiwara map out the timetable of centriole ciliation and deciliation in 3T3 fibroblasts and suggest that a correlation exists between the centriole cycle and the cycle of DNA replication and chromatid separation which is a more familiar feature of the cell cycle. Using indirect immunofluorescence, Tucker *et al.* confirm that centriole ciliation occurs when cells are arrested by conditions of low serum or high cell density in the post-mitotic phase of G_1 . When stimulated to divide by the addition of fresh serum, this unduplicated centriole deciliates within 1–2 hours, reciliates by 6–8 hours and undergoes a further deciliation between 12–24 hours which roughly coincides with the initiation of DNA synthesis. The microtubules of the mitotic spindle form part of an apparatus for dividing genetic material between sister cells and so it may not be surprising that microtubule and nucleic acid cycle are somehow coordinated. As these investigators point out, microtubule-

organising centres (for example the centriole and, probably, its associated material) seem to be diverted from alternative activities in preparation for mitosis at about the time that DNA replication is initiated. This is reminiscent of the sequence of events which occurs in the division cycle of the yeast *Saccharomyces cerevisiae* (Hartwell *Bact. Rev.* **38**, 164; 1974) where mutation of the gene controlling the replication of the spindle pole body (a microtubule-organising centre analogous to the centriolar complex) blocks the initiation of DNA synthesis and hence of other steps leading to mitosis. The timing of centriole duplication in mammalian cells is not yet determined, but may occur in late G₁ or early S (see also Rattner & Phillips *J. Cell Biol.* **57**, 359; 1973).

Although a useful marker for some cells, primary cilia are absent from others and it should not be forgotten that higher plant cells may not possess centrioles (see *News*

and Views **277**, 515; 1979). Higher plants do, however, possess amorphous microtubule-organising centres analogous to the amorphous type which surrounds the centriole in animal cells and plays a more direct part in forming the mitotic apparatus. Telzer and Rosenbaum (*J. Cell Biol.* **81**, 484; 1979) have isolated such pericentriolar material from HeLa cells and demonstrate that when isolated from mitotic cells it supports reassembly of microtubule protein into microtubules, but is ineffective when obtained from cells in interphase. Considered together, recent evidence points to a close association between the activity of microtubule-organising centres (both structured and amorphous) and the mitotic cycle. In view of this it may be useful to consider even the cells of multicellular organisms as 'differentiated flagellates' which undergo a metamorphosis to the reproductive stage driven by signals affecting microtubule organising centres. □

Previously, not all leukaemia patients could be shown to respond to their own leukaemia cells (Lee & Oliver *J. exp. Med.* **147**, 1334; 1978) but a combination of stimulation *in vitro* followed by expansion of the cytotoxic T cell population *in vitro* (up to 10 million-fold) reveals that virtually all patients can make an immune response to their own leukaemia cells. Furthermore, since stimulation with a pool of allogeneic lymphocytes was effective in generating kill against autologous leukaemia cells, it may be possible to generate cytotoxicity against a tumour even if no tumour material is available (important for immunotherapy of solid tumours where stimulator cells may be hard to obtain). That pooled allogeneic cells are effective reinforces the immunologists' dogma that transplantation antigens (HLA) in some way determine or are related to the T cell repertoire.

On the face of it the finding that T cells can be sensitised to autologous leukaemia cells and that other cells are not killed would seem to be strong evidence for tumour-specific antigens on leukaemia cells. However, this is a conclusion which should be regarded with considerable suspicion. Recent studies of leukaemia cells from many laboratories suggest that at least the majority of leukaemias have surface membrane phenotypes characteristic of normal, though sometimes rare, cells. A cautionary tale is provided by the acute lymphoblastic leukaemia (ALL) antigen which, although first discovered on leukaemia cells and thought to be ALL specific is in fact found on bone marrow stem cells in fetal and regenerating bone marrow (Greaves & Janosy *Biochim. biophys. Acta* **516**, 193; 1978). As yet no controls carried out in studies of autologous anti-leukaemia cytotoxic T cells have proved that the target antigen(s) for the killer cells are not just the same sort of normal differentiation antigen. If immunotherapy involved elimination of a stem cell population along with the leukaemic cells it might not be so attractive a proposition!

Delivery of immunotherapy is an additional problem. In animal experiments immunised cells have been most effective when administered mixed with tumour cells in local tumour growth assays. The relative inefficiency of systemically administered cells may be due either to incorrect migration *in vivo* of cultured lymphocytes or because more than one subset of lymphocytes is required for optimal elimination of tumour cells.

Even if there remain doubts about the use of T cells amplified *in vitro* for therapy, there can be no doubt that cultured T cells will be extremely useful tools. Not the least important of the questions which will surely soon be answered, doubtless using cloned sublines of killers, is what is the nature of human leukaemia antigens? If they turn out to be exclusively tumour-associated, the way to T cell immunotherapy may be open. □

Immunotherapeutic T cells?

from Peter Beverley

TWENTY years ago Thomas proposed the theory of immune surveillance. Since then it has been a hope of cancer immunologists that the immune system might be persuaded to do the work of physicians and surgeons and destroy the tumours of cancer patients.

Initially, because of their role in allograft rejection, thymus-derived lymphocytes (T cells) were thought to be the main agent of surveillance. Later data have thrown up other candidates for the role of 'surveyor' (Baldwin *Nature* **270**, 557; 1977) and the very idea of surveillance has been challenged (Moller & Moller *J. Natn. Cancer Inst.* **55**, 755; 1975). Nevertheless T cells remain an attractive possibility for immunotherapy because of their extreme specificity of response and because high levels of activity can be generated by stimulation *in vitro*.

The practicality of immunotherapy rests on a number of assumptions. First that tumours carry surface antigens distinct from those of normal cells; second that lymphoid cells of the tumour-bearing host can be immunised against the tumour antigens; and third that immune cells (or antibodies) when administered to the

tumour-bearing animal will be effective in reaching and destroying the tumour.

Experimental tumours often provoke an immune response, but in many cases it has been shown that immunity is directed against antigens coded by RNA tumour viruses carried in the tumour cells, while in general the evidence for immunogenic antigens on spontaneous tumours of animals is far weaker. In man apparently, tumour-associated antigens have been defined using heteroantisera but it is not clear whether these antigens provoke a T cell response by the tumour bearer. Similarly, although there have been many reports of cell-mediated immune responses to tumours in man, their exact nature and specificity have been difficult to define because the responses are weak and inconstant. Thus the two findings that T cells cytotoxic for human leukaemias could be generated (Sondel *et al. J. Immun.* **117**, 2197; 1976) and that specific cytotoxic cells could be propagated indefinitely *in vitro* (Gillis & Smith *Nature* **268**, 154; 1977) raised the hope that some of these problems might be resolved by the production of large numbers of highly active cytotoxic T cells directed at human tumour antigens.

This has now been achieved by Joyce Zarling and Fritz Bach, (this issue of *Nature*, page 685) and the power of the new methodology is certainly demonstrated.

Peter Beverley is an Imperial Cancer Research Fund Staff Scientist at the ICRF Human Tumour Immunology Group, University College Hospital Medical School, London.

Conserving genetic resources of animal species

The idea of the botanical gene bank, where seeds of rare and endangered plants are kept dormant in the deep freeze, to be revived when required, is now widely accepted. Here Dr B.N. Veprintsev of the Institute of Biological Physics and Dr Natalia Rott of the Institute of Developmental Biology of the USSR Academy of Sciences, argue that similar gene banks should be started immediately for rare and endangered animal species and speculate on how it might be possible in the future to reconstruct extinct animals from their germ cells stored in such a bank.

We are witnesses today to radical changes in the environment — the disappearance of virgin forest and prairie and the exploitation of hitherto undisturbed oceans, lakes and rivers. The disappearance of many animals and plants is an inevitable consequence of the loss of their natural habitats. Their loss would not only produce an aesthetically impoverished landscape but, more practically, eliminate important genetic resources and disrupt the stability of fragile ecosystems which have evolved over the centuries. The rate of change is accelerating rapidly. If we do nothing to save rare and endangered species now it will soon be too late.

Current efforts to preserve rare animals and plants in natural reservations, zoos and botanic gardens may not provide sufficient security for their future. We believe that we must find ways of conserving the basic genetic material of these species — the gonads, germ cells and somatic cells — so that they can be revived in the future when conditions for their survival may be more favourable.

We should like to propose various ways in which animals might eventually be reconstructed from the conserved genomes (Fig. 1). To put this hypothetical scheme into practice we need:

- to preserve the genetic information of endangered species in as complete a form as possible.
- to find ways of realising this information by obtaining zygotes and rearing animals from them.
- to find ways of changing the sex of some of the reconstructed animals if they are all of the same sex, as would inevitably follow from some of the methods proposed.

Sources of genetic information

The most realistic approach at present is to store frozen gonads, germ or somatic cells for subsequent reconstruction of zygotes. One immediate problem of course is how to obtain such material without further endangering the population of rare animals. In our opinion the most suitable cells for conservation would be the germ cells (oocytes and spermatozoa) and fertilised ova — the zygote itself before reduction cleavage.

For some species somatic cells might be an attractive genetic source as they are

easily collected and some at least could be expected to contain all the necessary information for development.

Storage

Deep freezing seems the best method of conservation at present. Already deep frozen sperm are widely used in animal breeding and more recently, methods of storing deep-frozen oocytes and embryos of cattle^{1,2} and some laboratory animals³ have been developed. Oocytes can be stored in gonadal tissue culture and in deep frozen gonads^{2,5} and subsequently can be matured either by implantation of the gonad into a female recipient⁵ or by hormone treatment *in vitro*^{4,6}.

Reconstruction of a viable zygote

The means of obtaining a viable zygote from the stored material must of course be the most speculative part of our scheme. In normal development, the interaction of the zygote nucleus with its cytoplasm is necessary for the proper expression of the genetic information. In differentiated somatic cells, however, although the nucleus contains a complete set of genetic instructions, most genes are not

expressed⁷. Therefore, in order to regenerate an animal from its conserved genome, it is necessary to provide an environment in which the genetic information is expressed as in normal development.

If we have only been able to conserve spermatozoa, one approach would be to fertilise the ova of a closely related species. The success of such interspecific hybridisation varies in different animals. In birds and fishes, hybrids between different genera and even families are possible. In mammals and amphibia interspecific hybridisation can take place. In cases where spermatozoa cannot fertilise the ova of another species this difficulty might be overcome by injecting the sperm, treated with proteolytic enzymes, directly into the ovum cytoplasm⁸. In some cases hybrid incompatibility might be overcome by obtaining tetraploid and amphiploid hybrids in which two chromosome sets belong to one species and two to the other. Such hybrids have been obtained in frogs and silkworms⁹. In all cases where the genetic material of two species is mingled, the required species would have to be re-established by repeated backcrossing of hybrid females with male sperm of the original species.

Another possibility might be androgenesis. This would avoid the problem of backcrossing to eliminate unwanted genetic material. Just as complete plants can be regenerated from male pollen alone, it might be possible to fertilise an ovum of another species, in which the nucleus had been inactivated, with the sperm of the desired animal. In this case one would obtain an, admittedly haploid, zygote, in which the genes were those of the required species but the supporting cytoplasm was of a different species. To get a diploid zygote two male

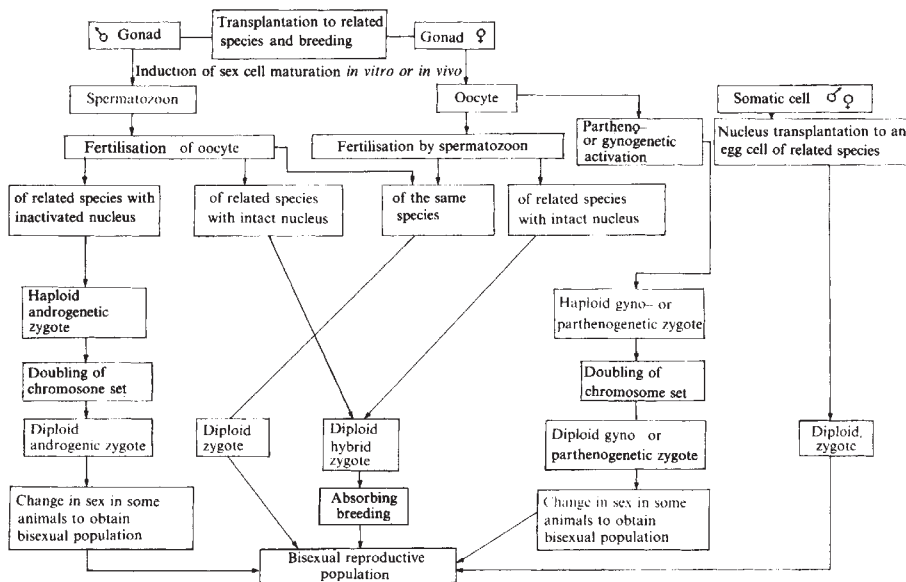


Fig.1

nuclei would have to fuse (dispermic androgenesis), and unfortunately so far this is known only in silk-worms¹⁰ and axolotls¹¹. Alternatively zygotes might be made diploid by inhibiting one of the primary cleavage divisions¹², but the resultant homozygous embryo might reveal lethal or undesirable recessive genes^{13,14}.

If only oocytes were available these might be induced to undergo parthenogenesis by chemical activation or by pseudo-fertilisation with sperm of a different species which cannot subsequently take part in development. A diploid embryo could again be obtained by inhibiting either one of the maturation divisions or one of the first cleavage divisions¹⁴; the problem of homozygosity remains.

If only somatic cells were available, perhaps the nuclei could be transplanted into the fertilised zygote of a closely related species, in which the nucleus had been inactivated. Experiments with amphibia of the same species have shown that in such cases normally inactive genes are reactivated and normal development proceeds. Microsurgical nuclear transplantation is successful in amphibia, fish and insects, and some limited success has been achieved with mammalian material¹⁵. Nuclei for one species can also be introduced into cells of another by cell fusion, brought about by Sendai virus or polyethylene glycol².

But nuclear transplation from one species to another has proved very difficult. In the many experiments which have been done, hybrids have been obtained in only four cases, and in two of them the oocyte nucleus remained in the recipient cell¹⁵. In most cases the embryos did not develop. Producing chimaeric embryos may help to overcome this difficulty. The cells of lethal hybrids were shown to have increased vitality when fused with the embryos of a normal parental species¹⁷. But although chimaeric

mice can be produced routinely by aggregating two very young embryos or by injecting cells from one into the other, attempts to obtain interspecific chimaeras present more problems.

Zygote development

How might reconstructed zygotes be reared to produce living animals? In insects, fish and amphibia this is not particularly difficult but for reptiles, birds and mammals the problems are enormous. No-one has yet achieved the complete development of such an embryo *in vitro*. In every case in which manipulated embryos have been successfully reared, they have been reimplanted into the mother or into a fostermother of the same species.

The most reasonable course in our scheme would be to transplant embryos 'conceived' *in vitro* into foster mothers of a closely-related species. However, attempts to do this — mouse into rat for example — have so far failed. The reasons for failure are unknown. There may be immunological rejection of the embryo; in this case it might be possible to induce tolerance. This would need suitable antigens from the 'preserved' species so we need to conserve not only gonads and germ cells but other tissue as well. For birds and reptiles interspecific ovarian

transplantation seems very attractive.

Sex determination

Some of the methods outlined above (parthenogenesis and gynogenesis for example) would produce animals all of the same sex, so it would be necessary to look for some way of converting some of them into the other sex. Mechanisms of sex determination differ from species to species so any treatment would have to be individually tailored. This may be less daunting than it seems, for there are species, for example among the fishes and amphibia¹⁶, that change sex spontaneously. And in other species hormones and other factors can influence sexual development.

Although most of the methods we have discussed are not yet practicable, research is proceeding at such a rate that they are beginning to look feasible in the not-too-distant future. Semen, blood and skin tissue from rare animals in zoos could be collected and stored without any danger to the animal itself. So we urge that collection and storage of suitable material should start now. But we need to be very cautious and not think that this approach can save the animal kingdom. It must be only one necessary step among many measures needed to meet the present emergency. □

The impact of climatic change

from Martin Parry

OUR knowledge of fluctuations of climate has exhibited a precocious flowering in the past 5 years and the time has come to consolidate the progress already made along many different paths and to point a way forward. This was the goal that 250 climatologists and historians from 33 countries set themselves recently at the conference on Climate and History at the Climatic Research Unit, University of East Anglia*.

The issues that emerged from this meeting touched on many facets of contemporary economics and society, in addition to the more specific problem of the impact of climate in history. The latter will be reported in detail elsewhere, but a number of points that were raised in papers and highlighted in discussions had far-reaching implications and deserve immediate debate. The points can broadly be divided into those that derive from a greatly increased understanding of the scale, direction and timing of climatic fluctuations, and those that are related to the economic and social impact of such fluctuations.

The reconstruction of past climates has

benefited from the recent advances made both in the quarrying of new data sources and the improved analysis of existing ones. At the University of Bern, Christian Pfister has recently developed a new set of procedures for standardising and storing a great variety of data, from weather descriptions in farming diaries to dates of lake freezing and the size of vine harvests. The system could provide the basis for an integrated and international pool of data — a pool to which information from new sources could continually be added. Several new data sources were mentioned at the conference, including tax records in Norway and meteorological records from China. At the same time, the improved 'calibration' of tree-ring data by Harold Fritts and his colleagues at the University of Tucson is giving us a vital bridge between the beginning of the instrumental records (about AD 1600 in Europe and AD 1800 in the New World) and the more distant record provided by studies of pollen, ice and sediments. In short, there is the exciting prospect of the construction of a climatic chronology of temperature and rainfall for much of the Northern Hemisphere over the last millennium that will be sufficiently accurate for us to evaluate the economic impact of the

1. Bilton R.J. & Moore N.W. *J. Reprod. Fert.* **50**,363 (1977).
2. Whittingham D.J. *Genetics* **78**,395 (1974).
3. Blandau R.J. & Odor D.L. in *Oogenesis* (ed. Biggers & Schultz) 496, University Park Press, Baltimore/Butterworths, London, 1974.
4. Steinberger A. & Steinberger E. in *Methods in Mammalian Embryology* (ed. Daniel) Freeman & Co, San Francisco, 1977.
5. Hafez E.E. *ibid.*
6. Erickson G.F. & Sorensen R.A. *J. exp. Zool.* **190**,123 (1974).
7. Gurdon J.B. *The Control of Gene Expression in Animal Development* Clarendon, Oxford, 1974.
8. Skobolina M.N. *J. Embryol. exp. Morph.* **36**,67 (1976).
9. Astaurov B.L. *Biol. Zbl.* **91**,137 (1972).
10. Astaurov B.L. & Ostrikova-Varshaver V.P. *J. Embryol. exp. Morph.* **5**,449 (1957).
11. Rott N. *Genetika* (USSR) **3**, (no 4) 44 (1965) (in Russian).
12. Hope P.C. & Illmensee K. *Proc. natn. Acad. Sci. U.S.A.* **74**,56 (1977).
13. Subtelny S. *J. exp. Zool.* **139**,263 (1958).
14. Beatty R.A. in *Fertilisation: Comparative Morphology, Biochemistry and Immunology* (eds Metz & Monroy) 1,413 Academic, New York/London, 1967.
15. Bromhall J.D. *Nature* **258**,719 (1975).
16. Gallien L. *Bull. Soc. Zool. France* **97**,233 (1972).
17. Baltzer F. *Rev. Suisse Zool.* **57**,93 (1950).
18. Gallien L. *Differentiation et organogenese sexuelle des metazoaires* Masson et Cie, Paris, 1973.

*Held on 8-14 July, 1979 and organised by the Climatic Research Unit.

fluctuations. As well as establishing our present climatic pattern quite firmly in the context of past changes, we can now say something about the range of alternative patterns of climate and weather we should be prepared for in the future. In particular, several papers reflected a significantly improved understanding of the nature of short-term climatic fluctuations, and the factors that distinguish them from long-term changes of the climatic pattern. The contributions from China were particularly welcome here because of the remarkably long and detailed run of records available.

The distinction between minor fluctuations and major changes may be arbitrary in terms of the mechanisms of climatic change, but it is important in comprehending their contrasting economic consequences. The ability of a rural economy, whether peasant or commercial, to absorb the shock of year-to-year changes in weather — or, especially, runs of adverse weather — is a critical factor in determining its economic success and its social harmony. Studies of climatic change and agricultural response strongly indicate the need for farming systems which will reduce the risks from uncertain weather by a spread of farming options. The irony is that improvements in technology and organisation enable us to respond more rapidly to drought in, for example, the Sahel or the US Great Plains; but, at the same time, they encourage farmers to adopt specialised and consequently risky farming systems that are more prone to suffer from climate changes. We can thus distinguish between the need for farming communities, especially in marginal areas, to be prepared and able to adapt to long-term changes of climate; and, on the other hand, the necessity of adopting farming systems which can absorb short-term climatic shocks without producing economic or social hardships.

These conclusions are drawn from several studies of marginal areas, where the nature of the climatic impact can more readily be interpreted. One of these is the cold marginal zone around the North Atlantic, whose response to climatic fluctuations has now been quite thoroughly researched. Analyses of options lost and gained by societies in Greenland, Iceland, Scotland and Norway were presented at the conference. The climatic constraint on agriculture in these areas is largely temperature; and, because this exhibits less variability than rainfall, any trend in its behaviour can be more readily detected, and suitable measures planned. The dry marginal zones in Africa, the Middle East and the US present a much greater problem, but there are encouraging signs of an increased understanding, particularly through the work of geographers based at Clark University,

Massachusetts.

Moreover, we are now beginning to comprehend the enormous social and economic cost of climatic fluctuations. There are now quite precise measures of the relationship between Indian food grain production and monsoon reliability; of changes in the US regional and national house heating bill resulting from the postwar decline in winter temperatures; and of correlations between wheat production and weather patterns on the Soviet steppes. The return on predicting climatic fluctuations is, of course enormous, but it is unfortunately still not in prospect. In the meantime, our increased knowledge of the nature of the fluctuations and their impact should enable us to plan an improved ability to absorb the shocks and to adapt to the changes. In this sense, the Climate and History Conference was a milestone, for it marked the move from the pioneer phase of initial chronology construction to a secondary phase of impact evaluation; and for the pioneers who are witnessing this coming of age — Hermann Flohn, Hubert Lamb and Gordon Manley — this must have been an extraordinarily satisfying experience. □

Life and death before birth

by Mary Lindley

LIFE before birth is fraught with hazards, and for man at least they often seem to be too much for the fetus, leading to spontaneous abortion and stillbirth, known collectively as fetal wastage. Some of the factors controlling the fate of the fetus were aired recently at the annual symposium of the European Society of Human Genetics*.

The timing of some of the events of fetal wastage is beginning to become clear from studies of developing mice. The fate of XO mouse embryos is being investigated by P. S. Burgoyne (University of Edinburgh). Using female mice heterozygous for a large inversion of the X chromosome, he has obtained XX, XO and XY zygotes from XX mothers. He has found that when XO embryos from such mothers are cultured, they are often abnormal by the morula/blastocyst stage. Up to 11 days of gestation they are normal, but immediately after implantation some are resorbed. Burgoyne suggested that the preimplantation abnormality leads to the postimplantation loss of XO embryos.

However, those embryos that survive the critical period seem to develop well, providing an intriguing parallel with the human situation. It has long been a puzzle

that XO human fetuses have a very high rate of spontaneous abortion in the first three months of gestation, while those that are not lost then will progress quite well by the standards of embryos with chromosome abnormalities. Burgoyne has also compared mouse embryos from the XX mothers possessing the inversion with embryos from XO mothers and found a greater degree of retardation in the latter. He suggests that some maternal effect influences development in addition to the chromosome abnormalities produced in the offspring.

A. Gropp and colleagues (Institute for Pathology, Lübeck) have now looked at all nineteen trisomies of the mouse, using the breeding programme they designed for the purpose. As would be expected from what is known of the fate of abnormal human embryos, different trisomies survive for different times. Three critical periods of development can be recognised. Some of the embryos, including trisomies for chromosome 17, do not survive a first phase of organogenesis up to days 11 and 12 of gestation. Days 13 and 14 constitute the second critical period and trisomic fetuses surviving this far may have gross malformations, for example of the heart. These include trisomies 12, 14 and 16. The third critical period, days 15 to 18, seems to be associated with impaired placental function and oxygen supply. Some trisomic mice survive birth, but their subsequent life span is limited by general hypoplasia.

The human fetus is of course much less convenient for study, and although fetal wastage is known to be extensive, its full extent can only be estimated. According to A. Boue (Centre for the Study of Prenatal Biology, Paris) the accepted dogma is that 15% of all conceptions end in recognisable fetal wastage, while 31% survive to birth. The other 54% are presumed to be lost very early in gestation, but they cannot be detected.

E. Williamson and J. Miller (Southampton General Hospital) have come up with a value of about 40% early post-implantation loss in their prospective study of 195 women who were trying to conceive. By taking regular measurements of human chorionic gonadotropin in the urine, Williamson and Miller were able to recognise when pregnancies began and then follow their course, some ending in early loss and others proceeding further. It has still not been possible to recognise the time of fertilisation, and figures of the sort produced in this study require a generous measure of speculation. It seems that there is no substitute for specimens, but at the very early stages of gestation the problems of collection are too great at present.

Some of the earliest hazards, preceding even conception, arise during meiosis. This

M.L. Parry is in the Department of Geography, University of Birmingham.

Mary Lindley is an Assistant Editor of Nature.

*The symposium, on Genetic Aspects of Fertility and Fetal Wastage, was held in Southampton on 19-21 July 1979 and organised by Dr. M. Seabright, Wessex Regional Cytogenetics Unit, Salisbury General Hospital. The abstracts will be published in the *Journal of Clinical Genetics*.

is still to a large extent a mysterious process, but clues to its dangers are beginning to emerge. A. G. Searle (MRC Radiobiology Unit, Harwell) pointed out that in male mice X-autosome translocations and a substantial number of autosomal reciprocal translocations (those with break points near one end of the chromosome) are particularly associated with a breakdown of meiosis, often due to failure of pairing. This leads to the impairment of spermatogenesis and almost complete male sterility. It is not clear why this should happen, and the puzzle is increased because similar translocations have no such drastic consequences in female mice.

During a session on sex determination A. Jost (College de France, Paris) pointed out some inadequacies in current thinking about the role of H-Y surface antigen in this process. The antigen has been widely discussed as the agent responsible for the differentiation of testes in mammals, in which the male is the heterogametic sex (XY), and of ovaries in birds and other vertebrates in which the female is the heterogametic sex (ZW). Jost pointed out that such a scheme is not compatible with the results of his own and earlier work on the development of freemartins. He said that the course of development is similar either when a female calf fetus is masculinised before birth in the presence of a male twin, or when a female axolotl or *Xenopus* is induced to become a male by having a testis grafted onto its ovary. In each case there is first a phase of no change, then profound retardation of ovarian development and finally, masculinisation, leading to complete sex reversal in the two amphibians. According to current theory, the role of H-Y in mammals is to masculinise rather than to suppress female organs, and Jost pointed out that this does not account for the early and prolonged ovarian inhibition. Also according to current theory, the heterogametic sex influences the homogametic sex, and this cannot explain why the heterogametic female amphibians were masculinised by the homogametic males. Some other explanatory scheme is needed, Jost said.

U. Wolf (University of Freiberg), the leading exponent of this work in Europe, explained his view of the role of H-Y antigen in development. From studies in his laboratory with gonads of rat and chicken, he concludes that H-Y acts as a signal inducing the undifferentiated gonad of the early embryo to become a testis in the XX/XY system or an ovary in the ZZ/ZW system. Thus the gonad begins to develop in the direction of one sex or the other according to whether H-Y is present or absent. Sex determination therefore involves control of the expression of a putative H-Y gene or genes specifying the production of H-Y antigen. Wolf believes that, although the H-Y gene is usually active only in the heterogametic sex, it is carried in the genome of both sexes. He

said that its presence in the homogametic sex is demonstrated when embryonic chicken testis is treated with oestrogen and converted to ovotestis which will then produce H-Y antigen. This may offer comfort to those who never believed that the H-Y gene was on the Y chromosome, and perhaps encourage rumours that it is really on the X chromosome.

More insights into sex determination and the role of sex chromosomes are promised by the Scandinavian wood lemming — billed as 'The incredible *Myopus*' and as usual, enthusiastically championed by K. Fredga (University of Lund). Wild populations of this small rodent (*Myopus schisticolor*) contain unexpectedly large proportions of females, some of which produce only daughters. Breeding data are bearing out the hypothesis that an X-linked mutation (designated X*) prevents the expression of masculinity in the presence of a Y chromosome, so that X*Y, as well as XX and X*X will be phenotypically female. Fredga said that the chromosomes X and X* can be distinguished by their

banding patterns. Most X*Y females produce only daughters because a unique form of non-disjunction gives rise only to eggs carrying the X* chromosome. A consequence of the failure of this mechanism is the production of an unusually high proportion of offspring with abnormal sex chromosome constitutions. Fredga and his colleagues in Lübeck, A. Gropp and H. Winking, have bred X*YY fertile females, X*XY sterile males and an X*XY true hermaphrodite with an ovary on one side of its body and a testis on the other. They believe this indicates that the normal X chromosome is necessary for male development, with at least two genes involved, possibly one structural and the other regulatory. Fredga suggested that the development of X*XY individuals is influenced by the pattern of X chromosome inactivation. Thus in the hermaphrodite, inactivation of the normal X would have resulted in the development of an ovary, while inactivation of X* would have resulted in a testis. Further news of *Myopus* will be awaited eagerly. □

Mars volatiles

from C. T. Pillinger

It is now almost 8 years since Mariner 9 went into orbit around Mars and found the planet shrouded by a gigantic dust storm to such an extent that only the very largest topographical features were visible. The storm abated and no doubt everyone remembers those tremendous first pictures of our nearest planetary neighbour. Later, in 1976, two Viking spacecraft, both soft landers and orbiters, continued the exploration of Mars. During the primary mission (June to November) little or no dust activity was observed. However, the scientific groups involved continued to monitor their instruments throughout the extended mission and data for an entire Martian year (approximately two Earth years) have been obtained. In that time 16 local and two global (although not as intense as the 1971 example) storms were observed. 'Mars Volatiles (and Meteorology and Dust)' (*J. Geophys. Res.* **84**, 136, 2793-3007; 1979) is a compilation of 16 papers by members of the Viking Volatile and Meteorology Working Group dealing with the results of the extended mission. Although various atmospheric problems are considered, dust and dust storms provide a common thread. As the working group ceased to have any official standing long ago, its members should be commended for the effort which must have gone into internal editing and cross-referencing so that readers are presented with a continuous story. All the papers are interesting and informative even to a casual reader looking for some general knowledge

about Mars. The complete volume or high quality reprints of any of the papers may be obtained from the American Geophysical Union.

It is not easy to summarise the cumulative efforts of such a diverse study but some of the points of most general interest might be as follows. (All references following are to the volume cited above). It appears that dust storms occur predominantly when Mars is near perihelion and originate mainly in the southern hemisphere possibly near the ice cap (Briggs *et al.*) Local storms are up to 8 km high and have abrupt tops (Peterfreund & Kieffer). Surface pressure oscillations indicate that thermally driven global atmospheric tides are involved with the major storms (Leory & Zirek). Perhaps not surprisingly, measurements of wind, surface and atmospheric temperature and pressure reveal that the arrival of storms is accompanied by significant changes in these parameters (Ryan *et al.*; Martin & Kieffer). Infrared thermal mapping may be able to distinguish ground fogs and water ice clouds from dust and to provide information about the structure of the lower atmosphere (Hunt). The observation that water vapour is uniformly distributed in the atmosphere explains the lack of predicted early morning diffuse fogs. Thus, water is not frozen out at night and

C. T. Pillinger is in the Department of Mineralogy and Petrology, University of Cambridge.



ICE on Mars again — This high-resolution photo of the surface of Mars was taken by Viking Lander 2 at its Utopia Planitia landing site on May 18, 1979, and relayed to Earth by Orbiter 1 on June 7. It shows a thin coating of water ice on the rocks and soil. The time the frost appeared corresponds almost exactly with the buildup of frost one Martian year (23 Earth months) ago. Then it remained on the surface for about 100 days. It is thought that dust particles in the atmosphere pick up bits of solid water. That combination is not heavy enough to settle to the ground. But carbon dioxide, which makes up 95% of the Martian atmosphere, freezes and adheres to the particles and they become heavy enough to sink. Warmed by the Sun, the surface evaporates the carbon dioxide and returns it to the atmosphere, leaving behind the water and dust. The ice seen in this picture, like that which formed one Martian year ago, is extremely thin, perhaps no more than one-thousandth of an inch thick.

Photo: NASA

exchanged between the surface and the atmosphere on a daily basis (Davies). The global and seasonal distribution of water vapour is consistent with a 10 cm to 1 m reservoir of H_2O in the regolith at latitudes above 40° . An asymmetrical transfer of water from south to north at a deposition rate of a few milligrams per square centimetre occurs possibly due to the propagation of storms in the south (Farmer and Doms). The spring-summer recession of the south pole observed by the Viking 2 orbiter was slower than previously reported by Mariner and Earth-based telescopes. Occurrence of a dust storm at the appropriate time might have been responsible (James *et al.*). Seasonal variations of up to 26% in atmospheric pressure are caused by evaporation of the ice caps — approximately 7.9×10^{12} tonnes or a 23 cm thick layer needs to be exchanged (Hess *et al.*). The temperature of the south polar cap in winter is significantly below the expected 148 K. One explanation is that there is a reduction of CO_2 in the atmosphere and a lowering of mean molecular weight near the surface (Hess). Wind speed and direction, pressure, temperature and optical depth measurements indicate the existence of low and high pressure 'weather' systems in the atmosphere (Tillman *et al.*). Optical depth studies also provide information about the shape and imaginary refractive index of the dust; the cross sectional weighted mean radius is estimated to be around $2.5 \mu m$. At the moment, water ice and CO_2 are being preferentially deposited in the north polar regions because of scavenging by the dust.

Equatorial and mid latitude regions are probably being eroded at a rate around 7 m per 10^6 yr; it may take 10^5 yr to produce the individual strata observable at the poles (Pollack *et al.*). Thus, cyclic variations in direction over geological time have led to sedimentary deposits of millions of cubic kilometres under both poles as testimony to the effects of dust storms. Parallel undulations of period a few kilometres

and height a few tens of metres observed near north polar surfaces may be the intermediate topographic features (Cutts *et al.*).

Perhaps the only unrelated, but nevertheless interesting paper (Carr), describes the formation of the chaotic terrain, thought to be Martian flood features, formed by the release of water from confined aquifers. □

New routes to polypeptide hormones

by Peter Newmark

THE production of polypeptide hormones has, until very recently, started either in the slaughter-house/mortuary or on the solid phase beads of the organic chemist. But at a recent meeting* it was clear that the knife and the spatula are yielding to the restriction enzyme, although not without a struggle.

A particularly elegant example of the bacterial route to the production of a human hormone was recounted by D. V. Goeddel (Genentech, San Francisco). Since the synthesis of the complete nucleotide sequence corresponding to the 191 amino acid human growth hormone

(hGH) would have been a monumental task and since the use of DNA copied from natural messenger RNA and then converted to double-stranded form (ds cDNA) would have coded for the precursor to hGH which bacteria would not be able to process to the mature hormone, the following approach was taken. First ds cDNA for hGH was cloned and treated with restriction enzymes to give a DNA sequence corresponding to all but the first 23 amino acids of the hormone. Then the synthetic polynucleotide corresponding to amino acids 1-23 and preceded by an ATG initiation codon was cloned. Finally the two fragments were ligated and cloned 'behind' a pair of *lac* promoters and a ribosome binding site.

*The 12th Annual Miles Symposium on Polypeptide Hormones was held in Baltimore, 11-13 July, 1979.

This led to a considerable yield of 186,000 copies of radioimmunoassayable hGH activity per *E. coli* cell ($2.4 \mu\text{g ml}^{-1}$). Furthermore, and as planned, the product was of about the correct molecular size, not being fused to a plasmid-determined protein as has been the case in most previous approaches to the bacterial production of hormones (see for example Martial *et al. Science* **205**, 602; 1979).

Given Genentech's efforts and their plans to market hGH in collaboration with Kabi, it was surprising to hear doubts expressed that more hGH is really needed. In the USA at least, said S. Raiti (National Pituitary Agency), although there were theoretical grounds for believing that several thousand untreated patients existed, in practice the medical demand for hGH was readily met by hGH isolated from cadaveric pituitaries. On the other hand he conceded that lashings of the hormone would allow proper trials of its mooted efficacy on burns, non-healing fractures, ulcers and haemopoiesis.

As yet, however, not even the biological activity of the presumed hGH has been tested. Indeed it remains to be shown that any bacterially-produced hormone has proper biological activity although Lydia Villa-Komaroff (University of Massachusetts, Worcester) reported that the substance released by trypsin treatment of fused β -lactamase-insulin produced by bacteria was somewhat active in a rat fat pad bioassay.

Another polypeptide (?hormone) that is attracting the attention of genetic manipulators is interferon. One only has to read *Omni* these days to discover that interferon is the great white hope of cancer therapy. The problem with both *Omni* and interferon is to separate science fiction from fact. For interferon that problem will only be overcome when it is available in sufficient quantities and in pure enough form for proper trials to be carried out. As yet the few, if encouraging, results that have emerged from clinical tests have relied on interferon that is less than 10% pure.

Greater purity is one of the main aims of both S. Pestka (Roche Institute of Molecular Biology, New Jersey) who is purifying interferon from chronic myeloid leukaemia serum and of C.B. Anfinsen (National Institutes of Health) who will soon be growing 1,000 litres of human lymphoblastoid cells as his source. Purity is important not only for clinical trials but for the determination of interferons' structure. So far it is thought to be a polypeptide of about 154 amino acids. The polypeptide is associated with carbohydrate components which, however, are not essential for activity. It is said (by others) that both at NIH and at Dupont the N-terminal amino acid sequence of interferon has been determined. If true that may give the genetic manipulators their entrée. For with

a synthetic polynucleotide complementary to the N-terminal amino acid sequence it may be possible to fish out pure messenger RNA for interferon to make cDNA for subsequent cloning.

Does the bacterial approach leave anything for the synthetic peptide chemist to do? Certainly! In the first place chemical synthesis still makes much sense for the smaller peptides, 'smaller' being less than 20 to 75 amino acids, depending on whose opinion one takes. It should however be noted that even for such a small peptide as the 14-amino acid somatostatin the bacterial route of production may be commercially viable; otherwise why will it become Genentech's first marketed chemical (for research purposes only) by the end of this year? Second, chemical synthesis is inherently 'cleaner' and has been considerably improved by the advent of continuous flow techniques with the prospect of real time monitoring of the products (B. W. Erickson, Rockefeller University). Third, chemical synthesis is well suited to the production of labelled peptides and peptide analogues. That is particularly the case when the analogues contain either modified or D-amino acids.

The synthetic peptide chemists have had, for example, considerable success in producing potent agonists and antagonists of the 10-amino acid peptide, luteinising-hormone-releasing-hormone (LHRH). Chiefly by the substitution of a D-amino acid at the sixth of the ten positions and by N-terminal changes, it has been possible to produce 'super LHRHs' with more than a hundredfold the potency of the natural hormone but also with paradoxical antifertility effects (see for example Ying & Guillemain, *Nature* **280**, 593; 1979). Considerable progress has also been made in producing analogues that are LHRH antagonists. The most active antagonists described by D. H. Coy (Tulane University, New Orleans) have parahalogenated D-phenylalanine at position 2 and acylated D-phenylalanine at position 1. Those substitutions together with a D-tryptophan at position 6 result in an antagonist that blocks ovulation at 0.03 mg per rat. Three years ago the best antagonists were 30-fold less potent.

The peptide chemists have had less success so far in improving on those thymic hormones whose structure has been determined, thymopoietin and FTS ('facteur thymique serique'). Thymopoietin is 49 amino acids long but a constituent pentapeptide has the biological activities of the full molecule (G. Goldstein, Ortho Pharmaceutical Co.). FTS is a nonapeptide first isolated from pig serum, now from human serum, and shown also to be present in thymic epithelium (J. F. Bach, Hôpital Necker, Paris).

Another serum peptide with effects on T cells is likely to be a tetrapeptide, possibly with a methylated amino acid at the N-terminus (A. Astaldi, Netherlands Red

Cross Blood Transfusion Service). Ironically the only clinical trials that were reported involved either the relatively crude 'thymosin fraction V' of A. W. Goldstein (George Washington University, Washington) or the unsequenced 'THF' polypeptide of N. Trainin (Weizmann Institute). Thymosin fraction V has recently proved effective in prolonging the survival of patients with oat cell carcinoma. And THF has induced a remarkable recovery in immunosuppressed children with lymphoproliferative diseases who had developed infections. The basis of these clinical successes, the cellular or molecular effects of thymic hormones and the relationship of the various contenders for that title remain mysterious.

The testing and production of polypeptide hormones is clearly a very competitive business. It will not be unprecedented if the clinical optimism evident at this stage has to be moderated when larger scale trials can be completed. Meanwhile amongst the genetic manipulators the perceived returns have led them to adopt Madison Avenue tactics. A sobering-up period will doubtless follow when the products have to be shown to be suited to clinical use. Meanwhile many claims are made to the effect that bacterial production is both cheaper and easier than present techniques. It will be interesting to see whether the eventual price as well as the availability of the hormones reflects current optimism. □



A hundred years ago

But there is, perhaps, nothing which shows more strikingly the identity of the protoplasm in plants and animals, and the absence of any deep-pervading difference between the life of the animal and that of the plant, than the fact that plants may be placed, just like animals, under the influence of anaesthetics.

We owe to Claude Bernard a series of interesting and most instructive experiments on the action of ether and chloroform on plants. He exposed to the vapour of ether a healthy and vigorous sensitive plant, by confining it under a bell-glass into which he introduced a sponge filled with ether. At the end of half an hour the plant was in a state of anaesthesia, all its leaflets remained fully extended, but they showed no tendency to shrink when touched. It was then withdrawn from the influence of the ether, when it gradually recovered its irritability, and finally responded, as before, to the touch.

It is obvious that the irritability of the protoplasm was here arrested by the anaesthetic, so that the plant became unable to give a response to the action of an external stimulus.

From *Nature* **20**, 21 August, 391; 1879.

review article

Nuclear vibrations

G. F. Bertsch

Michigan State University, East Lansing, Michigan 48824

Three types of nuclear vibrations have been observed in which a large fraction of the nucleons participate. There is a satisfactory explanation for the vibrational frequencies but not for the damping rates.

NUCLEI, like more familiar mechanical systems, can undergo simple vibratory motions. One mode of motion, in which protons move together in one direction while the neutrons simultaneously move in the opposite direction, has been known for 30 years. In the past decade, several new vibrational modes have been discovered. Our theoretical understanding of the nuclear motion has also considerably advanced. Although the most refined descriptions of the motion require a quantum many-body theory, the main features of the vibrations can be understood using classical physics. The motion of an extended body is described by a field of vectors which give the velocity of the medium at each point. Figure 1 shows the velocity fields for three types of vibration: monopole, dipole and quadrupole. These three modes are exhibited by nuclei as so-called 'giant vibrations', in which a large fraction of the nucleons participate in the motion. In both the monopole and quadrupole modes, all the nucleons move coherently in a direction which depends on their position in the nucleus. In the dipole motion, as mentioned above, protons and neutrons move in opposite directions.

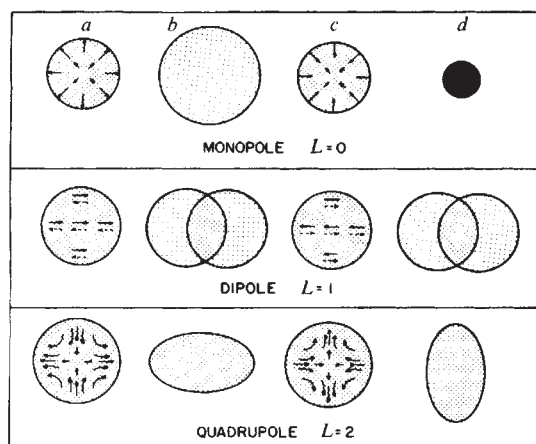


Fig. 1 The cycle of oscillation for simple vibrations. The velocity field at the moment when the nucleus passes through the equilibrium shape is shown in *a* and *c*. The shape at maximum distortion is shown in *b* and *d*. For the dipole, the neutrons and protons move in opposite directions. For other modes, the neutrons move with protons.

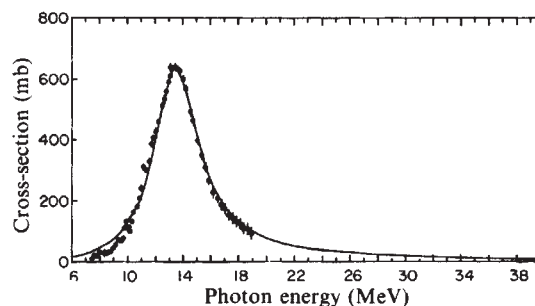


Fig. 2 The giant dipole resonance in ^{208}Pb , as observed by the photon reaction cross section (ref. 3, Fig. 32). The points are the experimental data, and the curve is the Lorentzian function.

Dipole vibration

The first nuclear vibration to be observed, the giant dipole, was found in measurements of photon absorption. The reaction cross-section is enhanced by resonance when the photon frequency matches the vibrational frequency. The first indications of the giant dipole resonance were obtained before photon beams with a broad energy range were available. Bothe and Gentner in 1937 had a source of 14 MeV photons and they measured the radioactivity produced in a variety of targets¹. The cross-sections varied considerably from target to target, indicating that there was a resonant absorption in some of the targets. By the mid-1940s, electron accelerators became available which produced bremsstrahlung photon beams of sufficient energy to observe the resonance. The improved data² showed that the absorption was concentrated in a resonance, the frequency of which varied systematically with the size of the nucleus.

One problem with these measurements is that the bremsstrahlung beam is broadly distributed in energy, so that the measurement for a particular energy required careful subtraction techniques. Currently, these experiments are done with photons that are produced by annihilation of positron, and are therefore monoenergetic. An example of the data now available³ with this improved technique is shown in Fig. 2. The points are experimental data, and the curve is the Lorentzian distribution,

$$\sigma(E) = \frac{\sigma_m}{1 + [(E - E_m)^2 / E^2 \Gamma^2]}$$

This formula has three parameters: a maximum strength σ_m , a centre frequency E_m , and a damping width Γ , which are adjusted to fit the data. For the heavier nuclei, the frequency depends on nuclear mass number A roughly as $\hbar\omega = 80/A^{1/3}$ MeV. The damping width is 4–5 MeV.

The spatial structure of the vibration, with protons moving against neutrons, is inferred indirectly from the data. Photons at the frequencies of interest have a long wavelength (~ 100 fm) on the scale of nuclear dimensions $R \sim 1.2 A^{1/3} \sim 3$ –8 fm. With a long wavelength, the force field is quite uniform and the photons cannot excite other kinds of motion besides the uniform displacement of protons against neutrons. Thus, as is the case with optical photons interacting with atoms, all but a negligible fraction of the cross-section is due to dipolar motion. It is useful to compare the magnitude of the cross-section with a theoretical limit known as a sum rule. The Thomas-Reiche-Kuhn sum rule for the absorption is given by⁴

$$\int_0^\infty \sigma(\omega) d\omega = \frac{2\pi^2 e^2}{Mc} \frac{NZ}{A}$$

where M is the mass of a nucleon, and N , Z are the neutron and proton numbers. The dipole vibration will satisfy this sum rule only if all the protons move coherently with a uniform translation. In the example shown in Fig. 2, the experimental absorption strength from the lorentzian fit is 3,000 MeV mb, which is just the limit from the sum rule. Thus, the motion in the dipole vibration seems to be completely coherent. This conclusion is modified by the presence of charged mesons.⁵

Quadrupole and monopole vibrations

As mentioned above, wavelengths shorter than can be provided by photons are needed to excite the monopole and quadrupole vibrations. The quantum relationship between momentum and wavelength, $p = h/\lambda$, implies that modes with more complex spatial structure can be excited only if enough momentum is given to the nucleus at the same time. This condition can be met by the inelastic scattering of massive projectiles. The first published experiments exhibiting a new vibration were done in 1957 with inelastic proton scattering at 180 MeV (ref. 6). Figure 3 shows these data; a distinct bump is visible at 18 MeV excitation. However, the significance of this bump was not realised until in 1971, when Bertrand and coworkers⁷ found the bump in their own data taken at lower energies and concluded that it was the quadrupole vibration of Fig. 1. Others also saw the vibration in electron scattering⁸, and offered a quadrupole vibration as one of the possible explanations.

The main evidence that the bump might be a quadrupole vibration was the angular distribution of the scattering cross-section. Each multipolarity has a characteristic shape to the angular distribution, which can be calculated by the Schrödinger equation for the projectile wavefunction. If the wavelength of the projectile is short compared to the dimensions of the target,

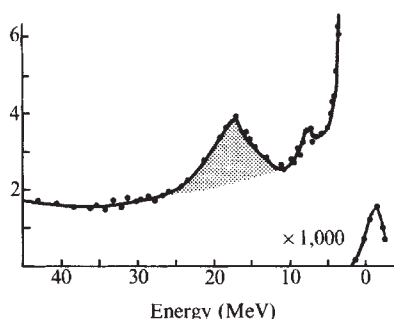


Fig. 3 Inelastic proton scattering spectrum for 180 MeV protons on a Ca target (ref. 5, Fig. 2). The stippled bump is the quadrupole vibration. The rise in cross-section towards zero excitation energy is the elastic scattering.

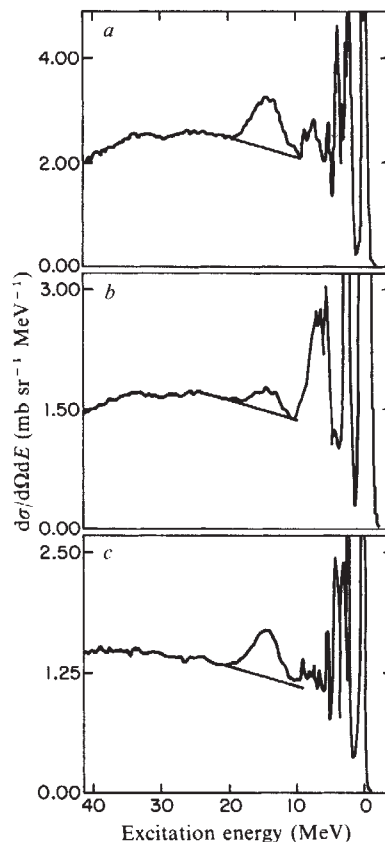


Fig. 4 Inelastic α -particle scattering on ^{90}Zr , taken at three different angles. *a*, $\theta_{\text{lab}} = 17.5^\circ$; *b*, $\theta_{\text{lab}} = 20.5^\circ$; *c*, $\theta_{\text{lab}} = 22.5^\circ$. The giant quadrupole is prominent at 17.5° and 22.5° , but is much reduced at 20.5° (ref. 9).

the angular distribution peaks at an angle given roughly by $\theta = L/\lambda R$, where L is the multipolarity of the excitation and R is the nuclear radius. At larger angles, the cross-section is oscillatory, with interference from the projectile passing opposite sides of the nucleus. Some recent data with angular dependence is shown in Fig. 4⁹.

The reaction is induced by α particles scattering on the nucleus ^{90}Zr . A prominent peak is seen at angles of 17.5° and 22.5° , but much reduced at 20.5° . This behaviour is only consistent with the predicted angular distributions for quadrupole or monopole vibrations. To decide between monopole and quadrupole, one must either look at smaller angles, or argue on the basis of the absolute strength of the peak.

As in photon absorption, sum rules can be used to analyse the absolute strength. However, the strength of the interaction for projectiles such as protons and α particles is not as well known *a priori* as it is for electromagnetic probes. To avoid this problem, the interaction between specific projectiles and target nucleons is adjusted by fitting the elastic scattering data. The interaction determined in this way describes inelastic transition quite well, and for many cases can be checked independently by its electromagnetic properties¹⁰.

Defining V as the potential field of the scattering projectile on the target, the inelastic scattering sum rule reads

$$\sum_f (E_f - E_i) \langle i | V | f \rangle^2 = \int d^3r n_0(r) \frac{(\vec{\nabla} V)^2}{2m}$$

where n_0 is the density of nucleons in the ground state and the excited states are labelled by f . This sum rule is easily understood in classical terms. The left-hand side is the average energy given to the nucleus by the collision. On the right-hand side, $\vec{\nabla} V(r)$ is the impulse given to a target nucleon at position r . Thus $(\vec{\nabla} V(r))^2/2m$ is the average kinetic energy given to a nucleon at r , and the integral is the average total energy given to

the nucleus. Since the integral involves only the ground state density, it can be reliably computed and compared with the actual strength of a given vibrational excitation. The peak in Fig. 4 has about 75% of the strength for a quadrupole vibration; if it were monopole the sum rule would be exceeded.

By now the systematic behaviour of the quadrupole vibration has been thoroughly studied¹¹. The mean frequency varies with mass number as $\omega \sim 63/A^{1/3}$ MeV/ $\hbar$. The width varies from 8 MeV in light nuclei to 3 MeV in heavy nuclei, with dips at magic nuclei.

In the past two years, experiments have been done at smaller angles, where the monopole and quadrupole have different behaviour. Inelastic scattering at 0° should favour the monopole vibration. The physics behind this is shown in Fig. 5 which shows a projectile wave enveloping a nucleus. The attractive interaction between the projectile and the target pulls the nucleus out along the equatorial region. This outwards motion can be resolved into a major monopole component and a minor quadrupole component. All the wavelets from the projectile will interfere coherently at 0° scattering angle. So, like the Poisson spot seen in optical diffraction around spheres, the scattering will show a pronounced peak at 0° .

The quantum calculations of the scattering support this picture: Fig. 6 shows calculations comparing the angular distributions of monopole and quadrupole at the most forward angles³⁶. These calculations treat the excitation of the vibration in the Born approximation, and use the solution of the Schrödinger equation in an optical potential for the projectile wavefunction. The sharp rise of the monopole near 0° is evident.

As the experiments have progressed to smaller angles with shorter wavelength projectiles, evidence has accumulated for a monopole vibration at a frequency slightly higher than the quadrupole. Figure 7 shows recent data (D. Youngblood, personal communication) for 0° inelastic α scattering. The centroid of the peak is at

$$\omega \sim 80/A^{1/3} \text{ (MeV}/\hbar)$$

Fortuitously, this is the same frequency as the dipole mode. The damping width of the vibration is about the same as for the quadrupole.

Theory

Paralleling the experimental developments, there have been considerable advances in the theory of the vibrations. The general theory, derived from many-body quantum mechanics in the 1950s, is known as RPA (random phase approximation). This theory was originally applied to nuclei^{12,13} in models which used an oversimplified description of the wavefunction. These models were instructive in exhibiting qualitative trends for the vibrational frequencies but they had severe deficiencies. The sum rules were not satisfied. The interaction between particles was unrealistic, and in many cases had to be adjusted to fit observed vibrational properties. In the past decade it has become possible to do the theory properly. The availability of

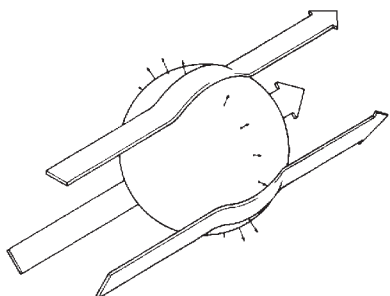


Fig. 5 A projectile wave, depicted with the broad arrows, envelops a nucleus and induces the transverse motion of the nucleus, shown with the small arrows. This motion resolves into a combination of monopole and quadrupole vibrations.

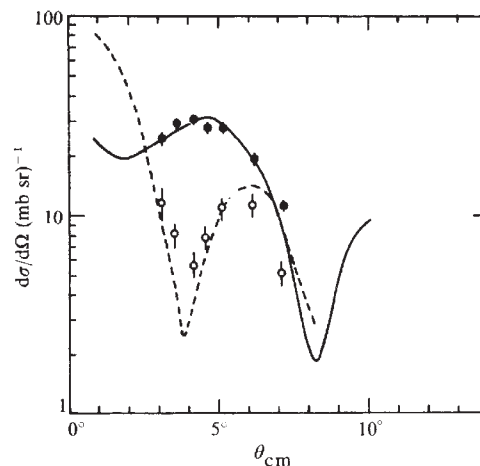


Fig. 6 Angular distribution for inelastic α scattering on ^{208}Pb (ref. 36). The solid line is the theoretical distribution for exciting a quadrupole vibration $E_x = 11.0$ MeV; $L = 2$, and the dashed line is the theoretical distribution for the monopole $E_x = 13.7$ MeV; $L = 0$. Points are data extracted by fitting the peaks like those in Fig. 4.

medium-sized computers means that the description of the wavefunction can be kept sufficiently general. In a configuration representation, the number of configurations for a heavy nucleus is typically several hundred. An example of an RPA calculation of the photon absorption strength is shown in Fig. 8¹⁴. The other key to a quantitative theory is the use of the same interaction hamiltonian to describe the vibrations as is used for the Hartree-Fock ground state. The important property of the interaction is that it contains a short range attraction which diminishes as the density increases. A simple parameterisation of the hamiltonian introduced by Skyrme has the form,

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{i < j} v(n_0, \mathbf{p}_i, \mathbf{p}_j) \delta(\mathbf{r}_i - \mathbf{r}_j)$$

where v is a polynomial function of density $n_0(r)$ and momentum $\mathbf{p}_i$. Such a hamiltonian can be determined by fitting ground-state sizes and binding energies¹⁶. More fundamental theories of the interaction, based on the free nucleon-nucleon potential, can be reduced to the Skyrme form.¹⁷ With the zero-range interaction, the Hartree-Fock ground states of nuclei are easily computed, as is the RPA response. With the largest computers, it is feasible to remove the restriction to zero-range interactions in the hamiltonian, and the most sophisticated RPA calculations are based on a more general parameterisation¹⁸. Using a hamiltonian consistent with ground state properties, the RPA gives an excellent description of the vibrations without any adjusted parameters¹⁹. The sum rules are automatically satisfied. The giant vibrations, with the exception of the monopole, are found at the proper frequency. The frequencies of low vibrations are generally reproduced to 10%. Figure 9 compares the theoretical²⁰ and experimental¹⁹ density fluctuation in a low-frequency vibration of ^{208}Pb . Note that the density fluctuation is strongest at a radius ~ 7 fm, which is at the nuclear surface.

The RPA can be simplified to algebraic formulae when the vibration exhausts the sum rule. These formulae are analogous to the classical formula for a vibrational frequency ω in terms of a mass M and spring constant k , $\omega^2 = k/M$. Before the RPA theory was known, classical models were proposed *ad hoc*. In the case of the dipole vibration, the mass is simply the mass of the protons and neutrons. The restoring force is due to the change in potential energy when the protons and neutrons are displaced with respect to each other. The dipole frequency was first calculated along these lines by Migdal in 1944 (ref. 22; see also refs 23, 24). The dipole motion in Fig. 1 physically separates neutrons from protons in the surface region: Migdal modified

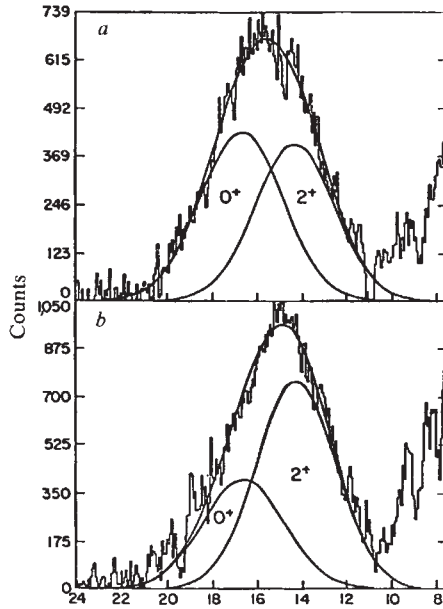


Fig. 7 Comparison of the α scattering at *a*, 0° ; and *b*, 4° . The peak is shifted upward in energy at 0° , due to the relative enhancement of the monopole at the forward angle.

the velocity fields to keep the particles together at the surface. The predicted frequency is then found to be

$$\omega \sim \left(\frac{4.3b_s}{MR^2} \right)^{1/2} \sim 80/A^{1/3} \text{ (MeV}/\hbar) \quad (1)$$

where $b_s \approx 50$ MeV is the symmetry energy of nuclear binding. The reduction of RPA yields a formula that contains additional factors arising from mesons and from a more accurate description of the velocity field^{25,26}. These corrections tend to cancel, and equation (1) is in excellent agreement with the empirical vibration frequencies of the heavier nuclei.

Migdal²² and others²⁷ also tried to calculate the quadrupole vibrational frequency, but in this case the classical ideas were simply wrong. It was assumed that the surface tension of the nuclear surface provided the restoring force to the spherical shape. The general formula for liquid drop vibrations was derived by Rayleigh, and is

$$\omega^2 = \frac{\sigma n_0}{MR^3} L(L-1)(L+2)$$

where σ is the surface tension. The formula can be evaluated using the nuclear surface tension derived from binding energy systematics. The predicted frequency is found to be ~ 2 MeV for heavy nuclei, which is much too low²⁷. However, as mentioned above, RPA calculations do predict the quadrupole vibration at the correct frequency. In fact, on the basis of such calculations Mottelson in the 1960s urged a search for the mode.

The reason why the classical calculation fails is that the main element in the restoring force is neglected. There is a resistance to shearing motion in any Fermi system caused by the complex spatial structure of the many-Fermion wavefunction. It turns out that the elastic shear modulus is proportional to the quantum kinetic energy of the particles. The proper formula for the vibrational frequency based on this restoring force is

$$\omega^2 = \frac{2\langle T \rangle}{M\langle r^2 \rangle} \quad (2)$$

Here $\langle r^2 \rangle$ is the mean square radius of the ground state, and $\langle T \rangle$ is the average kinetic energy per particle in the nucleus. Note that this formula, unlike the liquid drop formula, has the correct $A^{-1/3}$ dependence of frequency on mass number because $\langle r^2 \rangle \sim A^{2/3}$. When a simple estimate is made of the kinetic energy per

particle, for example, with the Fermi gas model or harmonic oscillator model, the predicted frequency from equation (2) agrees with the empirical model to better than 10%. As in the case of the dipole, equation (2) is somewhat oversimplified, since some of the subtler effects of the nuclear interactions are neglected²⁸.

We now consider the theory of the most recently discovered giant vibration, the monopole. The frequency of the monopole is especially interesting because it is directly related to the compressibility of nuclear matter. The compressibility is a fundamental property of the medium, but up to now there has been no way to measure it. There have been many calculations of nuclear compressibility beginning in the early 1940s, before the nuclear interaction was very well known. Techniques of calculations have since become very sophisticated, but we do not yet have a satisfactory understanding of the equilibrium density of nuclear matter²⁹. More detailed properties of the nuclear medium, such as its compressibility, are correspondingly more difficult to calculate. However, the existing calculations should serve as a guideline. Predicted values for the compressibility are in the range

$$k = V^2 \frac{\partial^2 E}{\partial V^2} = 2.5-5 \text{ (MeV fm}^{-3}\text{)}$$

with the higher number being more recent. The classical relationship between the radial vibration frequency and the compressibility of a liquid drop is $\omega^2 = \pi^2(k/Mn_0R^2)$. Applying this formula to finding k from the empirical ω , the compressibility of nuclear matter turns out to be $k \approx 3.5$ MeV fm⁻³. Having fixed the compressibility coefficient with the information from the monopole, it will now be possible to make more reliable calculations of other phenomena involving high-density nuclear matter. This includes high-energy heavy-ion collisions and the formation of neutron stars.

Damping

The damping of the nuclear vibrations is not well understood, theoretically or experimentally. There are two possible mechanisms for the damping. The vibration may disappear by emitting particles, which could be called evaporation; or the motion may mix with the more complex degrees of freedom in the many body system, which is analogous to friction. This complex state would eventually decay according to statistical considerations, which generally emphasise low-energy neutron emission. The measurement of decays of giant vibrations is a very active area of research. Some published experiments in the heavier nuclei claim that the decays are inconsistent with the statistical picture^{30,31}. On the theoretical side, the damping by particle emission is found to be unimportant except in light nuclei. In RPA, the damping rate due to nucleon emission can be

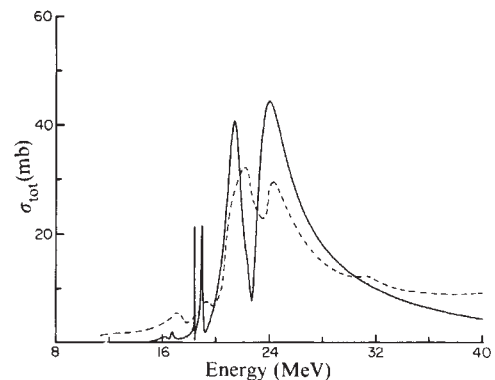


Fig. 8 The photon absorption strength of ^{16}O in the RPA theory (solid line)¹⁴, compared with experiment (dashed line). The resonance is split into five components by the shell structure, and has a width due to the decay by nucleon emission.

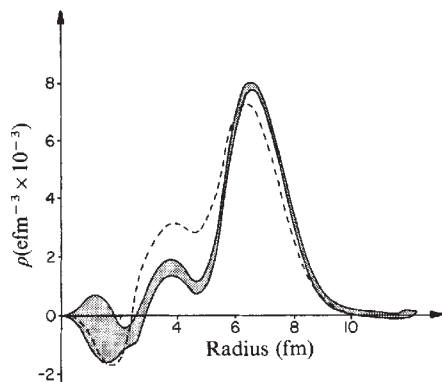


Fig. 9 The transition charge density $\langle 3^- | \rho | 0^+ \rangle$ for the lowest $L = 3$ state in ^{208}Pb . The experimental density²⁰ (hatched area) is measured by a Fourier-Bessel transform of the electron scattering form factor. The theoretical density¹⁹ is the result of an RPA calculation (dashed line), using a large basis for the wavefunction but a simple interaction.

calculated directly^{14,15}. The rate of emission of heavier particles, such as α particles, is more difficult to calculate but is also estimated to be small.

Nevertheless, the study of emitted particles has provided a useful tool for studying the vibrations. By examining a particular final transition, the emitted particle can be forced to carry away the entire angular momentum of the vibration. If a coordinate system is chosen in which there is no angular momentum about the z axis, the angular distribution of the particle then follows a Legendre function, $N(\theta) \propto (P_L(\theta))^2$.

Such a measurement was made for the giant quadrupole of ^{16}O , decaying by α -particle emission to the states in ^{12}C (ref. 32). The angular distribution for the ground state transition, which satisfies the above condition, is shown in Fig. 10. The data exhibit a peak at 0° of the right width to be described by $[P_2(\theta)]^2$. However, unlike the Legendre function, the data is not symmetric about 90° . This means there must be coherent mixtures of angular momenta. The fact that the background interferes with the decay of the vibration is not really unexpected. In the classical limit nucleon emission can take place by the projectile's knocking out a nucleon in the direction of the momentum transfer. The quantum description of this process requires interference from many angular momenta. It remains to be seen what theory predicts about how much interference is required.

The other form of damping is caused by mixing with more complicated modes of motion. A simple example of this is the behaviour of the dipole state in a deformed nucleus. A vibration in an arbitrary direction resolves into vibrations along the principal axes with different frequencies. This is seen in Fig. 11, showing the dipole strength for an ellipsoidal nucleus.

The lineshape can be well described as the sum of two Lorentzian curves. The lower frequency and higher frequency components correspond to the vibration along the major and minor axes respectively. The difference in frequencies agrees well with the R dependence in equation (1). Because there are two minor axes, the upper peak is predicted to have twice the strength which is found to be the case empirically. For the quadrupole vibration, there are three distinct frequencies instead of two, and the overall splitting is smaller because the motion has components along both principal axes. Consequently, for the quadrupole the deformation splitting is much more difficult to see.

The major source of damping is mixing with the statistical degrees of freedom of the many body system. In principle, this could be calculated from the shell model representation of nuclear wavefunctions, and several calculations along such lines have been reported^{33,34}. However, these calculations are not transparent, and it is more instructive to ask how the damping of the vibrations compares with damping of other kinds of motion.

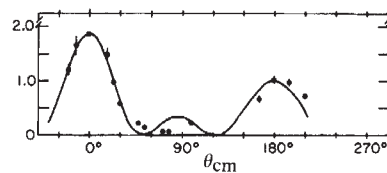


Fig. 10 Angular distribution of decay α particles following excitation of ^{16}O in the giant quadrupole region³². The 0° axis is along the momentum transfer direction. The probability distribution peaks in this direction, and also has a smaller peak in the opposite direction. $E_\alpha = 17.9\text{--}19.2$ MeV.

The simplest motion is that of a single nucleon across the nucleus. This is described quantum mechanically with an optical potential, and the finite lifetime of the nucleon τ is related to the imaginary part of the optical potential $\langle W \rangle$ by $\hbar/\tau = 2\langle W \rangle$. Optical potentials can be inferred from elastic scattering of low energy nucleons and from the capture of neutrons. The spreading of the single-particle strength can also be seen directly in the reactions in which a nucleon is transferred between the projectile and the target. These methods yield a damping width for nucleons close to the particle emission threshold of about $\hbar/\tau \approx 5$ MeV.

A naive model of the vibrational damping can then be constructed, utilising the representation of the wavefunction in terms of single-particle wavefunctions. The vibration is a superposition of one particle above the Fermi level together with a hole in the Fermi sea. Each of these should have a single particle lifetime so the total damping width would be twice that, hence $\hbar/\tau \approx 10$ MeV.

This estimate ignores a coherence in the vibrational motion: the fact that all of the particles are moving together inhibits the coupling to other degrees of freedom. This coherence is well-known in other quantum systems. For example, in metals the plasmon vibration might only be damped by 1 eV while an electron at the same energy has a damping many times larger. In fact, the detailed calculations of nuclear vibrational damping do show this quenching effect³⁵.

Although classical ideas are relevant to the description of the vibrational frequencies, they have not yet been helpful in discussions of damping. The classical concept for dissipation in extended systems is viscosity, and attempts have been made to apply this idea to the nuclear vibrations. Unfortunately, there is no fundamental justification for viscosity; it is only valid when particles have short mean free paths. In nuclei, the particles can travel completely across the nucleus with a high probability. The nuclear surface probably plays a more important part in the damping than do collisions between particles in the interior. A simple description of the damping in these circumstances is needed not only for the understanding of nuclear vibrations, but also for interpretation of data on heavy ion collisions.

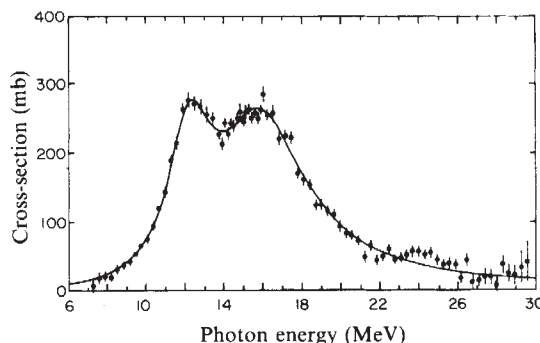


Fig. 11 The giant dipole resonance is ^{150}Gd , an ellipsoidal nucleus (ref. 3, Fig. 28). The curve is a sum of two Lorentzian functions, the upper having twice the area of the lower.

Prospects

What is the likelihood of discovering more vibrations? The possibility of more complex motion always exists, and other vibrations can be calculated in RPA. However, all the other modes of particle displacement have a higher mean frequency. The damping is predicted to be much larger, both within RPA and due to the coupling to other degrees of freedom. Thus it is unlikely that a coherent mode, containing most of the sum rule,

will be resolved as a distinct peak. In the inelastic scattering experiments, the spectrum contains a smooth background together with the presently known peaks, as may be seen in Figs 3 and 4. Part of this background at least must be due to highly damped vibrations of greater complexity. Without some specific measurable characteristic, however, it is not possible to identify them as such.

This research was supported by NSF grant PHY 7620097.

1. Bothe, W. & Gentner, W. *Z. Phys.* **106**, 236 (1937); **112**, 45 (1939).
2. Baldwin, G. & Klaiber, G. S. *Phys. Rev.* **73**, 1156 (1948).
3. Berman, B. L. & Fultz, S. C. *Rev. Mod. Phys.* **47**, 713 (1975).
4. Panofsky & Phillips, *Classical Electricity and Magnetism*, 335 (Addison-Wesley, New York, 1955).
5. Levinger, J. S. & Bethe, H. A. *Phys. Rev.* **78**, 115 (1950).
6. Tyren, H. & Maris, Th. A. *J. Nucl. Phys.* **4**, 637 (1957).
7. Lewis, M. B. & Bertand, F. E. *Nucl. Phys.* **A196**, 337 (1972).
8. Pitthan, R. & Walcher, T. *Phys. Lett.* **36B**, 563 (1971).
9. Youngblood, D. H. *et al. Phys. Rev.* **C13**, 994 (1976).
10. Bernstein, A. M. *Adv. Nuc. Phys.* **3**, 325 (1969).
11. Betrand, F. E. *A. Rev. Nucl. Sci.* **26**, 457 (1976).
12. Baranger, M. *Phys. Rev.* **120**, 957 (1960).
13. Brown, G., Evans, J. & Thouless, D. *Nucl. Phys.* **24**, 1 (1961).
14. Shlomo, S. & Bertsch, G. *Nucl. Phys.* **A243**, 507 (1975).
15. Krewald, S. *et al. Phys. Rev. Lett.* **33**, 1386 (1974).
16. Vautherin, D. & Brink, D. *Phys. Rev.* **C5**, 626 (1972).
17. Negele, J. & Vautherin, D. *Phys. Rev.* **C5**, 1472 (1972).
18. Blaizot, J. P. & Cogne, D. *Nucl. Phys.* **A284**, 429 (1977).
19. Bertsch, G. & Tsai, S. F. *Phys. Rep.* **18C**, 127 (1975).
20. Speth, J., Werner, E. & Wild, W. *Phys. Rep.* **33C**, 127 (1977).
21. Rothhaas, H. *et al. Phys. Lett.* **51B**, 23 (1974).
22. Migdal, A. J. *Phys. (Moscow)* **8**, 331 (1944).
23. Goldhaber, M. & Teller, E. *Phys. Rev.* **74**, 1046 (1948).
24. Steinwedel, H. & Jensen, J. A. D. *Z. Naturforsch.* **5A**, 413 (1950).
25. Myers, W. & Swiatecki, W. *Phys. Rev.* **C15**, 2032 (1977).
26. Bertsch, G. & Stricker, K. *Phys. Rev.* **C13**, 1312 (1976).
27. Bohr, A. & Mottelson, B. *Kgl. Dan. Vid. Selsk. Med.* **27**, No. 16 (1953).
28. Golin, M. & Zamick, L. *Nucl. Phys.* **A249**, 320 (1979).
29. Day, B. D. *Rev. Mod. Phys.* **50**, 495 (1978).
30. Wolynec, E., Dodge, W. R. & Hayward, E. *Phys. Rev. Lett.* **42**, 27 (1979).
31. van der Plicht, J. *et al. Phys. Rev. Lett.* **42**, 1121 (1979).
32. Knöpfle, K. T. *et al. Phys. Lett.* **74B**, 191 (1978).
33. Danos, M. & Greiner, W. *Phys. Rev.* **B138**, 876 (1965).
34. Soloviev, V. G., Stoyanov, C. & Vdovin, A. *Nucl. Phys.* **A288**, 376 (1977).
35. Bertsch, G. *et al. Phys. Lett.* **80B**, 161 (1979).
36. Youngblood, D. H. *et al. Phys. Rev. Lett.* **39**, 1188 (1977).

articles

A 30,000-yr isotope climatic record from Antarctic ice

C. Lorius*, L. Merlivat†, J. Jouzel† & M. Pourchet*

* Laboratoire de Glaciologie du CNRS, 2, rue Tres Cloîtres 38031 Grenoble Cedex, France

† Laboratoire de Géochimie de l'Eau, DRA Saclay, B.P. 2, 91190 Gif/Yvette, France

Simple glaciological conditions at Dome C in east Antarctica have made possible a more detailed and accurate interpretation of an ice core to 950 m depth spanning some 32,000 yr than that obtained from earlier ice cores. Dated events in comparable marine core has enabled the reduction of accumulation rate during the last ice age to be estimated. Climatic events recorded in the ice core indicate that the warmest Holocene period in the Southern Hemisphere occurred at an earlier date than in the Northern Hemisphere.

ALTHOUGH the stable isotopic composition ($^{18}\text{O}/^{16}\text{O}$, D/H) of polar ice sheets can provide a continuous record of past climatic conditions, the interpretation of the isotope profiles in terms of climatic temperature changes over a time scale ranging back to the last ice age is complicated by several factors including: (1) the determination of the time scale; (2) the influence of the elevation at which the snow was deposited and of ice sheet

stability; (3) the establishment of a transfer function from isotopic δ ratio to temperature in past conditions. Despite these limitations the study of ice cores has already provided very useful palaeoclimatic data¹⁻⁵. Similar limitations occur in the interpretation of other indirect sources of climatic data.

As isotopic records from areas near an ice divide are simpler to interpret, thermal core drilling to 906 m depth was carried out at Dome C (74° 39' S; 124° 10' E; elevation: 3,240 m, mean annual temperature: -53.5°C, Fig. 1) during the 1977-78 Antarctic field season as part of the International Antarctic Glaciological project⁶.

The isotope profile

The core recovery was about 98%. Sampling was carried out in the field by cutting a continuous slice from along the length of cleaned ice core. The $^{18}\text{O}/^{16}\text{O}$ ratios were measured with a fully automatic double mass spectrometer with an accuracy of 0.15‰ in the δ scale relative to the V. SMOW. The results averaged for samples of about 4 m length are plotted in Fig. 2 with depths expressed in metres of ice equivalent to take account of the

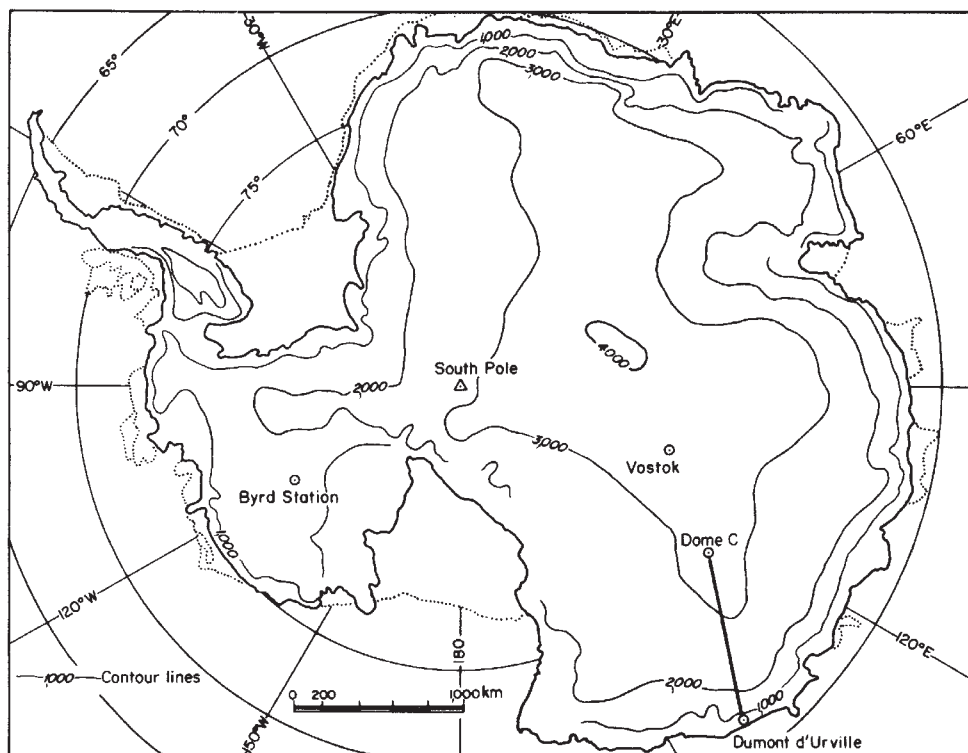


Fig. 1 Map of Antarctica showing the location of several stations (including the drilling sites) and of the traverse from Dumont d'Urville towards Dome C (dashed line). Recent data³² indicate a dome-like shape around the Dome C site.

lower densities in the upper firn layers. Each increment samples a time interval a little longer than one century. A smooth curve was drawn from adjustment on the basis of a spline function of the order of 2 with a parameter of adjustment $\rho = 0.001$. The main features of the δ curve indicate relatively uniform conditions down to 381 m, with a mean value of -49.9% (standard deviation $\sigma = 0.46$) followed by a rapid decrease down to 510 m. From then and down to 687 m δ fluctuates around -55.3% ($\sigma = 0.37$); isotopic fluctuations and presumably those of climate are of similar magnitude in this depth interval to those of the upper 381 m, when mean values over the order of 100 yr are considered. The general trend is then towards less negative values down to the bottom (875 m, ice equivalent). Significant events are superimposed on these main features; some of these are indicated in Fig. 2.

Ice core chronology

Different methods may be used for dating deep ice cores^{7,8} including measurements of radioactive decay, periodic changes of various parameters, and specific horizons dated from other sources. Some of these studies will be performed later, although they may not be successful in the low accumulation of the Dome C area. At present, a preliminary time scale can be calculated from the simplest ice flow model⁹ which assumes a uniform vertical strain rate and a constant value of snow accumulation. Although this model has other limitations¹⁰ it can be applied with reasonable confidence to ice dome areas when bedrock is flat and when considering the upper part of the ice sheet; both these conditions are fulfilled for the Dome C ice core. However, a large uncertainty results from the assumption of steady-state conditions over a time range which goes back to the last ice age. This is particularly true for the rate of snow accumulation which has so far been determined only over the past 25 yr from the radioactive layers formed from nuclear tests ($0.037 \text{ m ice yr}^{-1} \pm 0.004$). The age of the ice (t yr) with depth can be calculated from

$$t = \frac{H}{\lambda} \ln \frac{H}{H-z} \quad (1)$$

where λ is the accumulation rate, H the total ice thickness

(3,400 m) and z the depth, all expressed in metres of ice equivalent. The ages so calculated are given in Fig. 2 (numbers in parentheses).

If the main isotopic features reflect climatic changes, we can try to get chronological information from climatic records obtained from other sources. Data from continental areas of the Southern Hemisphere are scarce but some reference data may be obtained from marine sediment cores which provide a climatic record of surface seawater temperature, assuming that typical climatic events are synchronous with those recorded in Antarctic snows. Previous results¹¹ from core MD 73025 (Indian Ocean; $43^\circ 49' \text{S}$, $51^\circ 19' \text{E}$) have been provided by J. C. Duplessy and G. Delibrias (personal communication). Eight ^{14}C ages were determined on the top 3.5 m of this core and the comparison of pelagic and benthic foraminifera isotopic profiles shows that some of the characteristic events recorded in the Dome C ice core are also apparent in this marine core. The selected climatic features can be seen in Fig. 2; depth in both cores with the three ^{14}C interpolated marine sediments ages are given in Table 1. According to Duplessy these climatic features can be recognised in several northern and southern ocean deep-sea cores having a high accumulation rate with roughly the same ^{14}C ages within experimental errors.

Table 1 Depth and interpolated marine sediment age of Dome C and MD 73025 cores

Dome C ice core Depth (m ice equivalent)	MD 73025 marine core Depth (m)	^{14}C (yr)
381	2.36 ± 0.02	$10,550 \pm 300$
437	2.82 ± 0.02	$12,600 \pm 300$
510	3.10 ± 0.05	$15,500 \pm 500$

Snow accumulation rate changes

From these marine dates, we calculate by equation (1) a mean rate of snow accumulation in different ^{14}C time intervals. This gives 0.038 m yr^{-1} for the past 10,550 yr, 0.031 m yr^{-1} from 10,550 to 12,600 yr BP and 0.029 m yr^{-1} from 12,600 to 15,500 yr BP.

Although there are many uncertainties in such an approach and the accuracy of the accumulation figures may not be better than 10–15% (Table 1), the two last values are based on differences between ^{14}C ages, thus reducing problems of a possible time lag between marine and ice records and of errors of ^{14}C dating.

The rate of accumulation over the past 10,000 yr agrees very well with the value (0.037) measured over the past 25 yr. Since we know there are short-term changes in the accumulation rate this may seem to be a coincidence. However, the conclusion that present-day accumulation rates are typical of those for the past few thousand years has already been stated¹² for both the Greenland and Antarctic ice sheets from dating⁷ and inferred from observed temperature–depth gradients¹³. Our figures suggest a decrease of the rate of snow accumulation associated with colder climatic conditions, the accumulation 14,000 yr ago being about 75% of the present value.

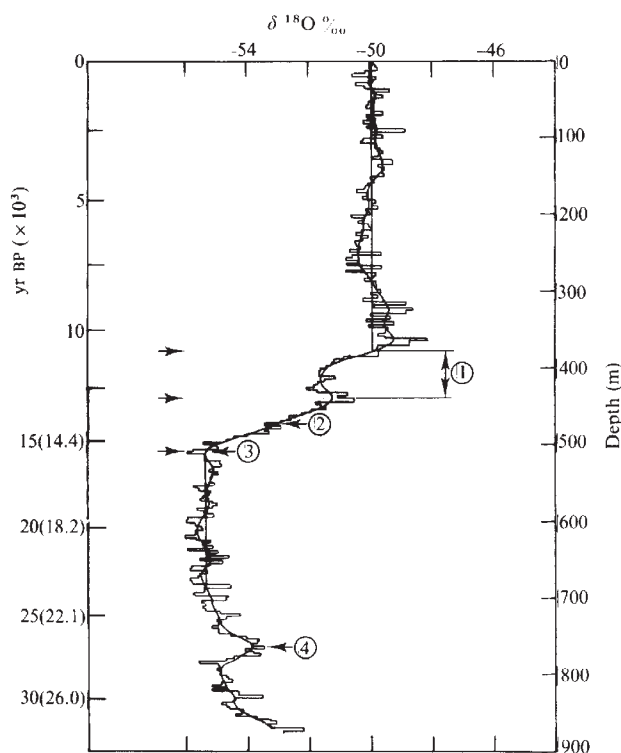


Fig. 2 δ record from the Dome C ice core plotted for about 4-m long increments and smoothed curve. The straight lines indicate mean values over the past 11,000 yr and during the late part of the last ice age. Significant isotopic events are numbered on the right hand side. On the left hand side arrows indicate reference levels corresponding to ^{14}C dates obtained from marine sediments. All depths are expressed in metres of ice equivalent. Ages are estimated from a simple ice flow model assuming a variable or constant (numbers in brackets) rate of accumulation.

Note that the mean annual accumulation in Antarctica is primarily governed by the amount of water vapour that can be carried in the atmosphere circulating above the surface inversion layer¹². This is governed by the atmospheric temperature and by the saturation water vapour pressure over ice above the inversion. For condensation temperatures of around -40°C in the Dome C area, a ratio of 0.75 for the vapour pressures corresponds to a condensation temperature change of about 3°C .

The study of precipitation¹⁴ at the South Pole gives a gradient of $d\delta^{18}\text{O}/dT_c = 1.4\%$ per $^\circ\text{C}$, where T_c is the condensation temperature obtained from radiosonde measurements of cloud

temperature. Analysis of snow samples collected between Dumont d'Urville and Dome C (Fig. 1) has provided an empirical relationship¹⁵ between the mean surface temperature T_m and the isotopic deuterium content of $\delta\text{D}\text{‰} = 6.04 T_m (^\circ\text{C}) - 51$ equivalent to $d\delta^{18}\text{O}/dT_m = 0.75\%$ per $^\circ\text{C}$. In Antarctica mean temperatures above the atmospheric inversion layer (T_i) are similar to the condensation temperatures¹², while the inversion strength is linearly related¹⁶ to mean surface temperature. A survey of existing data for eight stations with T_m values in the range from -10 to -56°C gives $dT_i/dT_m = 0.67$ with a correlation coefficient of 0.99. Assuming $T_i = T_c$ leads us to interpret our traverse results of equivalent to $d\delta^{18}\text{O}/dT_c = 1.1\%$ per $^\circ\text{C}$, a value which is in reasonable agreement with that obtained at the South Pole.

A change of 3°C for the temperature of condensation previously deduced from accumulation data would lead to a change of 3.3% between ice recently deposited and that deposited 14,000 yr ago. The agreement with the value obtained from the Dome C record (3.1%) is very satisfactory considering the large number of assumptions involved, which confirms conclusions about snow accumulation changes inferred from the Indian Ocean marine core data.

To obtain a continuous time scale we use relation (1) with a constant rate of accumulation ($0.037 \text{ m ice yr}^{-1}$) down to a depth of 381 m, and then decreasing to 75% of this value down to 510 m; this last figure is used down to 687 m. Further down it does not make a significant difference whether or not we assume an increasing accumulation linked to the $\delta^{18}\text{O}\text{‰}$ change. The calculated time scale is given in Fig. 2. According to this model the age of the bottom of the 905 m deep ice core is about 32,000 yr.

Climatic interpretation

Two other ice cores with depths comparable or greater than that from Dome C have been recovered in Antarctica (Fig. 1). At Byrd Station ($80^\circ 01' \text{ S}$, $119^\circ 31' \text{ W}$) the 2,160-m deep ice core covers a much longer time period and the isotopic change towards much more negative values is located between 1,000 and 1,400 m. The location and characteristics of the Vostok ice core are more comparable to those of Dome C one, the isotopic change at Vostok occurring between 250 and 375 m. Part of the published^{2,17} isotope profiles plotted against ages are shown in Fig. 3.

Comparison of Figs 2 and 3 shows similar trends in the three δ profiles. Event 1 from Fig. 2 is well marked in all three locations and there is a parallelism between the Byrd and Dome C isotopic profiles down to event 3. Events 1, 2 and 3 are also apparent in the Vostok ice core when looking at a more detailed isotope profile⁴, although there is a large scattering of the δ values probably due to discontinuous sampling technique. The less negative values of event 4 are also observed in the other ice cores.

Dating problems make precise age comparisons difficult, but it must be pointed out that the depths of the four events in the Vostok and Byrd ice cores relative to that of the Dome C record show a rather constant value. For the Vostok ice core where $\lambda = 0.024 \text{ m ice yr}^{-1}$ this value is similar to the ratio of the present rates of snow accumulation; this is not the case for the Byrd ice core, probably due to the very complicated regime in this area.

It appears then that the main characteristics of the δ profiles are similar for east Antarctica (Dome C, Vostok) and west Antarctica. Thus they probably reflect climatic events which affected both east and west Antarctica, rather than any ice sheet instability¹⁸ because these two distinct areas have very different glaciological regimes. Furthermore, isotopic events observed in the Dome C ice core are also seen in the record of sea-surface temperature changes in marine sediments.

The ice core chronology from Dome C indicates a relatively cold period from 11,000 to 13,000 yr BP; this is shown also in the record of past temperature of the New Guinea Highland sum-

marised in ref. 19 from careful pollen analysis. The Dome C fluctuations observed during this period could be equivalent to the European Dryas-Bölling-Alleröd stages depicted in the Camp Century (Greenland) ice core¹ although dating uncertainties prevent any conclusion about their synchronism.

The last major glacial peak occurs around 16,000 yr BP, which agrees with the much less detailed pollen record from New Guinea¹⁹ and the indication from various sources²⁰⁻²² that continental warming had probably begun in the Australian-New Guinea area by 16,000 $\pm$ 2,000 BP. This worldwide climatic feature has been found in many marine sediment cores both in the Northern and Southern Hemispheres^{23,24} with an age of about 18,000 yr which may vary locally by several thousand years.

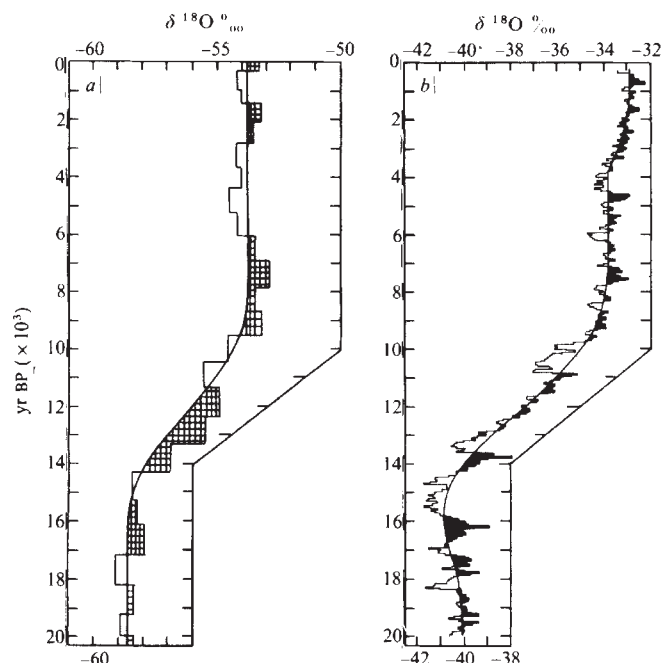


Fig. 3 The Vostok (a) and Byrd (b) δ records during the past 20,000 yr from ref. 2 (Byrd) and ref. 17 (Vostok). The δ values exceeding the smoothed curves are shown in black.

The warmer conditions shown on the Dome C δ record before 25,000 yr BP also agrees with conditions observed around this time in New Guinea by pollen analysis²² and with North America glacial chronology²⁵.

The isotopic shift associated with the transition from the coldest part of the ice age towards mean climatic conditions from the last 10,000 yr is 5‰ for Vostok¹, 7‰ for Byrd² and 5.4‰ for Dome C. When values for Vostok and Byrd station are corrected¹² to allow for the origin further inland of the deeper ice and for a higher surface elevation of the ice sheet during the last ice age, the $\delta^{18}\text{O}$ ‰ change associated with climatic changes would become 4.5 (Vostok) and 5.3 (Byrd). Although corrections for the same factors at Dome C could be made if measurements of total gas content²⁶ were available, severe cracking in the ice core from Dome C has made such a study impracticable. Since Dome C is at present a centre of outflow and if we assume that the form of the ice sheet has changed little over the past 10,000 yr, then the δ values in the upper 381 m of ice core should provide a useful climatic record over this period.

However, at the end of the last ice age, a lower sea level may have resulted in a slightly expanded ice sheet over east Antarctica, leading to greater ice thicknesses and increased surface elevations. These effects would be much greater in coastal areas, while in central east Antarctica estimates of the consequent surface rise are around 100 m (refs 4, 28). Such a rise, together

with the present δ -elevation gradient observed in the Dome C area¹⁵ would involve a correction of 1.5‰ to allow for changed elevation during the ice age. Thus it seems that the climatic change at the end of the last ice age was similar at Dome C, Vostok and Byrd, but further studies of ice thickness changes at Dome C are needed to confirm this conclusion.

The isotopic shift has still to be corrected for modifications of oceanic parameters during the last ice age. The most obvious is the change of mean isotopic value of oceanic waters. Due to the storage of water in δ -depleted ice sheets an increase of 1.6 $\delta^{18}\text{O}$ ‰ with reference to the present conditions has been estimated¹¹. This effect will probably be reflected in antarctic precipitation but from previous discussion it could be compensated by greater ice thickness during the last ice age.

The isotopic composition of ice cores, although principally governed by the temperature of condensation, can be affected by many other parameters¹. However, it has been suggested¹² that the present empirical relationship between mean surface temperature and isotopic content may not have been too different in Antarctica during the last ice age. As previously mentioned a 0.75‰ per °C $\delta^{18}\text{O}$ change was obtained in this area for a large range of temperatures, which leads to a tentative estimation of a difference of about 7 °C for the surface temperature between the coldest part of the ice age and the present climate. From previous discussions the suggested changes for the temperatures of condensation and above the inversion layer are reduced by about half.

Although there are many assumptions in such interpretations, they agree with the results of atmospheric modelling of the ice age (18,000 yr BP) climate^{29,30} which also indicate lower precipitation during cold conditions comparable with those described here. For the past 11,000 yr the climatic record inferred from the δ profiles (Fig. 2) shows that the warmest conditions occurred during the period 11,000–8,000 yr BP and the coldest conditions from about 8,000 to 4,000 yr ago. For the past 4,000 yr we observe a moderately warm period followed by a slight trend towards less positive δ values, with present surface conditions close to the mean value for Holocene.

Figure 3 shows that the Vostok δ profile indicates the same succession of climatic periods with a slightly different chronology. However, the Byrd record shows no apparent agreement with this sequence² due to the difficulty of interpreting its complicated glaciological regime.

The lack of comparative data from the Southern Hemisphere makes it difficult to estimate the spatial representativity of the Dome C Holocene record. During this period, the climatic evolution inferred from our δ profile gives a different picture from that obtained from New Guinea pollen analysis^{19,20}. The warmest period (11,000–8,000 yr BP) that we observed at the end of the transition from the last ice age is much earlier than the usual estimates for the so-called hypsithermal in the Northern Hemisphere. This is, however, in excellent agreement with the fact that the warmest sea surface temperature since the last glacial age³¹ in the Southern Hemisphere have a radiocarbon date of 9,400 yr (24).

This work was supported by Terres Australes et Antarctiques Françaises, NSF (Division of Polar Programs), DGRST and INAG. We thank all participants in drilling, laboratory and computation work, and J. C. Duplessy, G. Delibrias and G. de Q. Robin for helpful discussions.

Received 2 May; accepted 1 July 1979.

1. Dansgaard, W., Johnsen, S. J., Clausen, H. B. & Langway, C. C. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 37–56 (Yale University Press, 1971).
2. Johnsen, S. J., Dansgaard, W., Clausen, H. B. & Langway, C. C. *Nature* **235**, 429–434, and 236, 249 (1972).
3. Epstein, S., Sharp, R. P. & Gow, A. J. *Science* **168**, 1570–1572 (1970).
4. Barkov, N. I., Gordinenko, F. G., Korotkevich, E. S. & Kotlyakov, V. M. in *Proc. IUGG Symp.*, 382–387 (IAHS Pub. 118, 1977).
5. Paterson, W. S. B. *et al.*, *Nature* **266**, 508–511 (1977).
6. Lorius, C. & Donnou, D. *Courrier de CNRS* **30**, 6–17 (1978).
7. Hammer, C. U. *et al.*, *Glaciology* **20**, 3–26 (1978).
8. Lorius, C. in *Isotopic and Temperature Profiles in Ice Sheets* (ed. Robin, G. de Q.) (Cambridge University Press, in the press).
9. Nye, J. F. *Proc. R. Soc. A* **239**, 113–133 (1957).

10. Lliboutry, L. *Traité de Glaciologie*, Vol. 2, 578 (Masson et Cie, Paris, 1965).
11. Duplessy, J. C. in *Climatic Change* (ed. Gribbin, J.) 46–67 (Cambridge University Press, 1978).
12. Robin, G. de Q. *Phil. Trans. R. Soc. B* **280**, 143–168 (1977).
13. Budd, W. F., Jessen, D. & Radok, U. *Nature* **232**, 84–85 (1971).
14. Aldaz, L. & Deutsch, S. *Earth planet. Sci. Lett.* **3**, 267–274 (1967).
15. Lorius, C. & Merlivat, L. in *Proc. IUGG Symp.* 127–137 (IAHS Publ. 118, 1977).
16. Phillpot, H. R. & Zillman, J. W. *J. geophys. Res.* **75**, 4161–4169 (1970).
17. Barkov, N. I., Gordienko, F. G., Korotkevich, E. S. & Kotlyakov, V. M. *Inf. Bull. Sov. Ant. Exp.* **90**, 39–49 (1975).
18. Jessen, D. in *Climatic Change and Variability, A Southern Perspective* (eds Pittcock, A. B., Frakes, L. A., Jessen, D., Paterson, J. A. & Zillman, J. W.) 77–81 (Cambridge University Press, 1978).
19. Webster, P. J. & Streten, N. A. *Quat. Res.* **10**, 279–309 (1978).
20. Dodson, J. *Quat. Res.* **8**, 97–114 (1977).
21. Walker, D. in *Climatic Change and Variability, A Southern Perspective* (eds Pittcock, A. B., Frakes, L. A., Jessen, D., Paterson, J. A. & Zillman, J. W.) 82–97 (Cambridge University Press, 1978).
22. Bowler, J. M., Hope, G. S., Jennings, J. N., Singh, G. & Walker, D. *Quat. Res.* **6**, 359–394 (1976).
23. Climap project members *Science* **191**, 1131–1137 (1976).
24. Shackleton, N. J. in *Climatic Change and Variability, A Southern Perspective* (eds Pittcock, A. B., Frakes, L. A., Jessen, D., Paterson, J. A. & Zillman, J. W.) 73–76 (Cambridge University Press, 1978).
25. Goldthwait, R. P., Dreimanis, A., Forsyth, J. L., Karrow, P. F. & White, G. W. in *The Quaternary of the United States* (eds Wright Jr, and Frey, D. G.) 737 (Princeton University Press, 1965).
26. Raynaud, D. & Lorius, C. *Nature* **243**, 283–284 (1973).
27. Budd, W. F. & Radok, U. *Rep. Prog. Phys.*, **34**, 1–70 (1971).
28. Schilling, D. & Hughes, T. *Int. Symp. Dynamics of Large Ice Masses*, Ottawa, 1978 (in the press).
29. Gates, W. L. *Science* **191**, 1138–1144 (1976).
30. Manabe, S. & Hahn, D. G. *J. geophys. Res.* **82**, 3889–3911 (1977).
31. Hays, J. D., Lozano, J., Shackleton, N. & Irving, G. in *Investigation of Late Quaternary Paleooceanography and Paleoclimatology* (eds Cline, R. M. & Hays, J. D.) 337–372 (Geological Society of American Memoir, 145, 1976).
32. Scott Polar Research Institute, Cambridge, Map: Antarctica (Radio echo sounding map series A. 1:5000000. Sheet 3, 1974).

Shelf sea fronts' adjustments revealed by satellite IR imagery

J. H. Simpson & D. Bowers

Marine Sciences Laboratories, University College of North Wales, Menai Bridge, Anglesey, UK

IR imagery from NOAA 5 has been used to study the movement of shelf sea fronts. After the elimination of tidal advection, the fronts show remarkable consistency of position and do not adjust to the strong semi-monthly variation in tidal stirring which indicates the operation of a feedback process in vertical mixing.

IR IMAGERY from NOAA 4/5 has provided convincing evidence of the role of tidal mixing in controlling the distribution of sea surface temperatures in the European shelf seas. In particular it has been shown that the positions of the surface manifestations of the shelf sea fronts are closely parallel to contours of the h/u^3 parameter^{1,2}, as previously predicted³.

As the archive of imagery grows, it should be possible to use it to answer more detailed questions about the evolution of the fronts and the processes occurring on them. Indications have already been obtained from several images of the development of large scale (~25 km) instabilities in the along front flow which are clearly important as agencies of cross-frontal mixing^{1,4}.

In this article we attempt a more systematic analysis of two years data from NOAA 5 to investigate the extent to which the mean positions of the fronts adjust to variations in both the rate of tidal mixing and the surface heat flux.

Equilibrium adjustment

We particularly wish to examine the possibility that the mean frontal positions vary during the semi-monthly range cycle of the tides. As the amplitude of the tidal streams increases by a factor ~1.8 from neaps to springs, the intensity of tidal dissipation, which is proportional to u^3 , will rise by a factor ~5.8. The balance between surface heating and tidal and wind stirring at a front may be expressed in a constant efficiency model¹ as

$$R = 1 - a \left(\frac{\bar{u}^3}{Qh} \right) - b \left(\frac{\bar{w}^3}{Qh} \right) = 0$$

where h is the depth; u the tidal stream velocity; w the wind-speed; Q the surface heating rate, and the parameters a and b ,

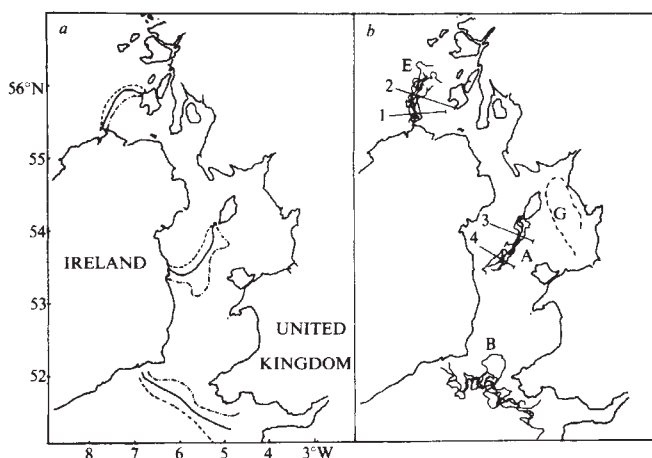


Fig. 1 *a*, Contours of $\log_{10} h/u^3$ --- 1.5; — 1.9; --- 2.2. h is the depth (m) and u_s is the tidal stream amplitude at springs ($m s^{-1}$). The separation of the contours 1.5 and 2.2 indicates the range of hypothetical adjustment between neaps and springs. Contour shape based on Pingree and Griffiths². *b*, Composite of all frontal positions observed in July 1977 from NOAA 5 imagery. Fronts are labelled as follows: A, Western Irish Sea (WIS); B, Celtic Sea; E, Islay; G, Liverpool Bay. Reference lines, along which displacements were measured, are numbered.

which contain the efficiencies of tide and wind mixing, may be assumed constant. The overbar denotes a time average.

If we further assume that Q and w^3 are spatially uniform over a particular frontal region, the 'equilibrium positions' should be defined by

$$\chi = \frac{h}{u^3} = \text{constant}$$

and, for slow changes in $\bar{u}^3$, the front would move to the appropriate contour of χ . This hypothetical adjustment, for a range change corresponding to the phase inequality cycle is shown in Fig. 1*a*. In the case of the Islay front, for example, the equilibrium displacement would be ~20 km. In other frontal

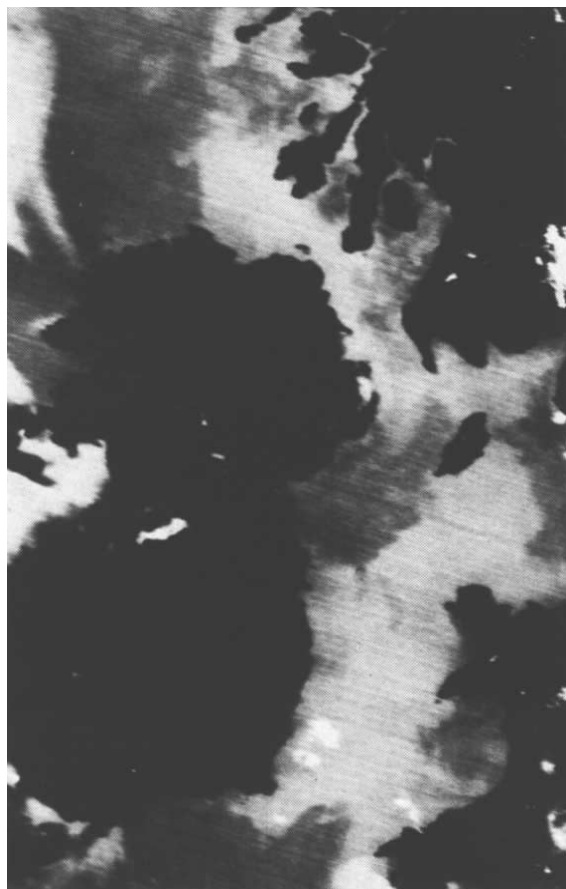


Fig. 2 NOAA 5 IR image from 21 May 1978, at 09.30 GMT, showing Islay and WIS fronts.

regions this movement would be generally larger due to the weaker gradients in χ .

On the other hand, if the time scale for adjustment is long in relation to the fortnightly mixing cycle, adjustment may be substantially reduced. The mixing model of James⁵, which incorporates a vertical transfer rate dependent on Richardson number, and hence stability, predicts negligible movement. Over longer periods we might also anticipate that the mean position of the front would adjust to changes in the heating rate Q which, at latitude 55°N , falls from a maximum of $\sim 140\text{ J m}^{-2}\text{ s}^{-1}$ in mid-summer (day 172) to $\sim 30\text{ J m}^{-2}\text{ s}^{-1}$ at 31 August (day 243) and reaches zero around mid September. Less regular adjustments due to variations in the level of wind mixing are also a possible source of variability in frontal positions.

Data acquisition and processing

The analogue data from the $10.5\text{--}12.5\text{ }\mu\text{m}$ band radiometer on NOAA 5 are received at the University of Dundee where they are converted to imagery with appropriate grey scales by a modified photo-facsimile machine. In the example shown in Fig. 2, where the darkness of the image increases in proportion to the radiation temperature, the positions of the Islay and western Irish Sea fronts are clearly visible. The 'front' is here taken to be the point of maximum surface gradient in sea surface temperature as judged from grey scale changes on the analogue images. Direct observations at sea indicate that it is close to, though not exactly coincident with, the transition to complete vertical mixing.

The photographs are corrected, by analogue techniques, for geometric distortion due to slant angle viewing of the Earth's surface from the spacecraft. As we are using the imagery to detect relatively small displacements of geographical position, it is important to quantify the extent of any residual distortion in the imagery. We have estimated this by determining, from a series of images, the relative positions of two headlands whose separation is known. The standard deviations of 17 measurements was $\sigma = 1.4\text{ km}$ for a separation of 61 km , a value which is representative of the distance from coastal reference points at which fronts are located. With increasing distance from the coast the relative uncertainty of positions will be larger, while close to land it is limited by the spatial resolution of the scanner which is $\sim 1\text{ km}$.

For each cloud-free picture of a frontal region the displacement of the front along lines approximately perpendicular to the mean frontal direction was measured from a datum marker. Figure 1b shows the location of the reference lines together with a superposition of tracings of the frontal positions as seen on all the useful images during July 1977. The number of cloud-free views shown here is typical of the data yield for summer months. Other fronts, notably one evident in the original pictures in Liverpool Bay (Fig. 2) will be described elsewhere.

An indication of the variability of frontal positions may, of course, be obtained from composite diagrams like Fig. 1b but note that the apparent displacement here includes a contribution from tidal advection, by the stream component normal to the front, which may be as great as $\pm 8\text{ km}$. To remove this, we considered the attractive possibility of using the known times of sampling, which defines the relevant phase of the tide, together with the observed positions of the fronts to extract the tidal streams by regression analysis. This approach is feasible and an example is given below but there is an inherent danger. As the satellite is Sun-synchronous the sample time is constrained to essentially the same time each day. For the UK the morning pass occurs on average at $\sim 10.00\text{ GMT}$. Consequently the principal solar constituent S_2 in the tidal signal is sampled at an almost constant phase while M_2 is observed with a phase slowly changing at a rate of $24.38^\circ\text{ d}^{-1}$. This generates an apparent displacement with a semi-monthly periodicity as can be seen in Fig.

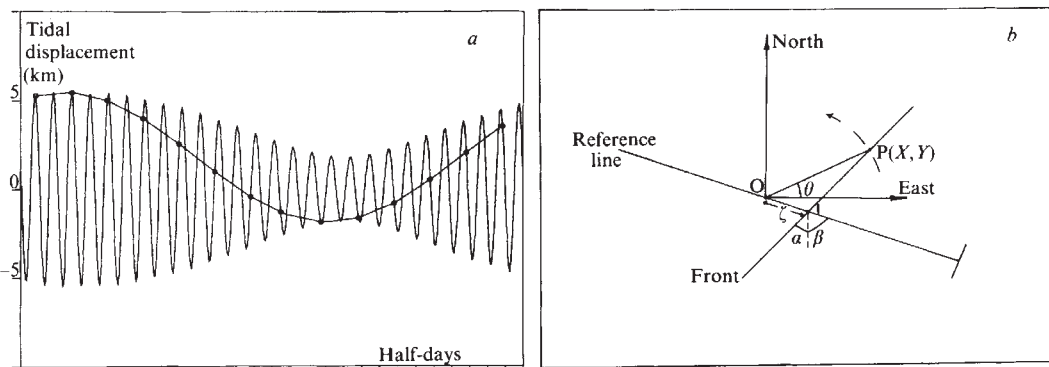


Fig. 3 *a*, Predicted tidal displacements for Islay line 2 based on current meter determination of M_2 and S_2 . Circles indicate the effect of sampling daily at a fixed solar time of 10.00 GMT. The result is a signal of semi-monthly periodicity which may be ambiguous with springs-neaps adjustment. *b*, Displacement of the intersection point I of the front and reference line by an elliptical tidal excursion of a water particle P about its mean position O.

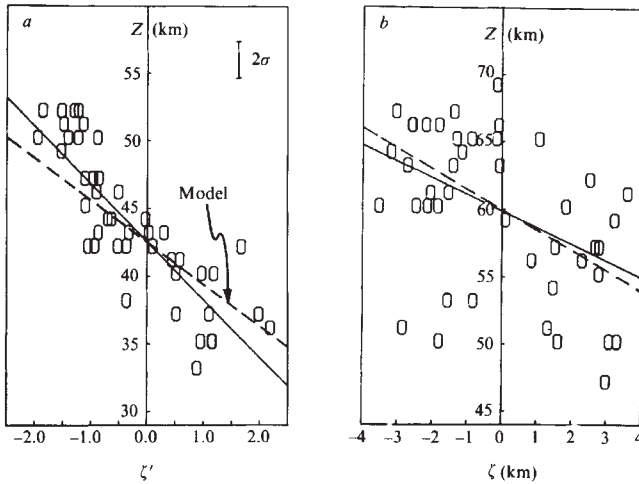


Fig. 4 Displacement Z versus a , predicted tidal displacement ζ' for Islay line 1, and b , predicted tidal displacement ζ for Islay line 2. The continuous lines are the results of a least squares fit. The dashed lines are based on a numerical model.

3a. This displacement will be indistinguishable from the neaps-springs mixing adjustment if the phases coincide. For a two component tide it can be shown that the phase difference of the range cycle and the apparent displacement is just

$$\Delta\nu = \phi_s - \omega_m \tau$$

where ϕ_s is the phase lead of the solar tide, ω_m is the frequency of the lunar tide and τ is the time before midday at which sampling takes place.

Ideally the ambiguity should be removed by doubling the sampling frequency using data from two satellites. In practice this is hardly necessary because of the spread in sampling time (± 1.5 h) about the mean and the fact that the evening pass occurs on average 10 h (not 12 h) after the morning transit, an asymmetry which results from the inclination of the satellite orbit to the Earth's axis.

If an adequate sample of velocity data is available from a mooring near the front, prediction of tidal displacements X and Y of a water particle may be made by the usual harmonic method according to

$$X = \sum_{i=1}^N \frac{a_i}{\omega_i} \sin(\omega_i t + \kappa_i) \quad \text{and} \quad Y = \sum_{i=1}^N \frac{b_i}{\omega_i} \sin(\omega_i t + \lambda_i)$$

where the a_i , b_i , κ_i , λ_i are the amplitudes and phases determined by least squares harmonic analysis of the velocity data. The displacement of the front along the reference line, which is not the same as the displacement of a water particle, may then be found from

$$\zeta = \frac{(X^2 + Y^2)^{1/2} \cos(\theta + \alpha)}{\sin(\alpha + \beta)} \quad \text{where} \quad \theta = \tan^{-1} Y/X$$

where the angles are defined in Fig. 3b. This approach assumes that the front is straight for a length large compared with the section of it which crosses the reference line during a tidal cycle. This requirement is least severe if we choose β so that $\alpha + \beta = \pi/2$, when the range of ζ has a minimum.

Where adequate tidal stream data is not available, we have compared the observed displacements to a non-dimensional tidal displacement defined by

$$\zeta' = F(t) \sin(\omega t + \phi)$$

where ω is the semi-diurnal tidal frequency and ϕ is a phase lead determined by reference to tidal atlas data. The slowly varying factor $F(t)$ is defined by

$$F(t) = \frac{\text{Elevation range at time } t}{\text{Elevation range at mean neaps}}$$

and found from elevation predictions at a convenient port. Comparison of ζ and ζ' in a case where suitable data were available showed acceptable agreement with an r.m.s. difference of 1.1 km.

Results

The Islay front: The tidal advection signal is illustrated in Fig. 4 where the observed displacements Z on lines 1 and 2 have been plotted against ζ' (Fig. 4a) and ζ (Fig. 4b). In the latter case, current meter data were available from a closely adjacent current meter station⁶. The slope of the regression line is 1.22 with a standard deviation of slope $\sigma_m = 0.44$, so that the observed displacements are consistent with the predictions although the scatter is considerable.

A better fit is apparent in Fig. 4a in which a regression line is compared with an estimate based on the numerical model of Pingree and Griffiths². The difference in slope (1.1 km) is significant at the 1% level, but this may partly be due to the interpolation of values between model grid points in a region of high gradients. The observed slope difference corresponds to a change in position of only one grid interval of 9 km.

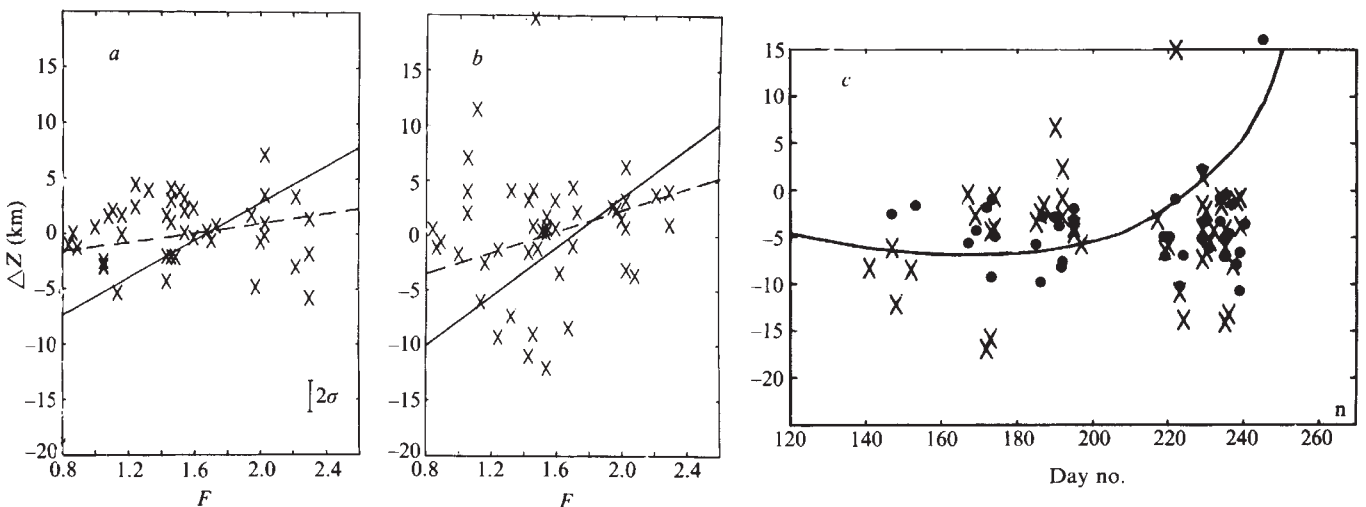


Fig. 5 *a* and *b*, residual displacement ΔZ against non-dimensional range index F for *a* Islay line 1; *b* Islay line 2. The extent of equilibrium adjustment is shown by the solid line. The dashed line has the largest slope consistent with the data points at 95% confidence ($F = 1$ at mean neaps). *c*, Residual displacement ΔZ versus day number n for the Islay front: $\bullet$, line 1; $\times$, line 2. The curve shows the form of the equilibrium adjustment to the seasonally variable heat flux.

An alternative explanation is the distortion of the Z - ζ relationship by curvature of the front, which we have neglected. When sufficient data are available, systematic departures from a straight line may be used to estimate frontal curvature and distinguish between fixed meanders and evolving eddies whose contributions should be random.

Subtracting the tidal advection signal, we obtain estimates of the residual displacement ΔZ which are plotted against tidal range in Fig. 5 along with a line which represents the extent of equilibrium adjustment. The data points are clearly inconsistent with such adjustment and the slope of the regression line through them is not significantly different from zero. The most probable adjustment between neaps ($F = 1$) and springs ($F = 2$), is determined from the regression as 0.25 km for line 1 and 1.2 km for line 2. Upper bounds are 1.9 and 4.8 km respectively at the 95% confidence level.

The marked difference in scatter between the two lines is apparently due to the front's tendency to form eddies which increases northwards. The lack of such instabilities on the Islay front near the Irish coast contrasts with that commonly observed on the shelf sea fronts, and may mean that observations of adjustment here are untypical of these features. Moreover the Islay front exhibits a strong salinity contract which distinguishes

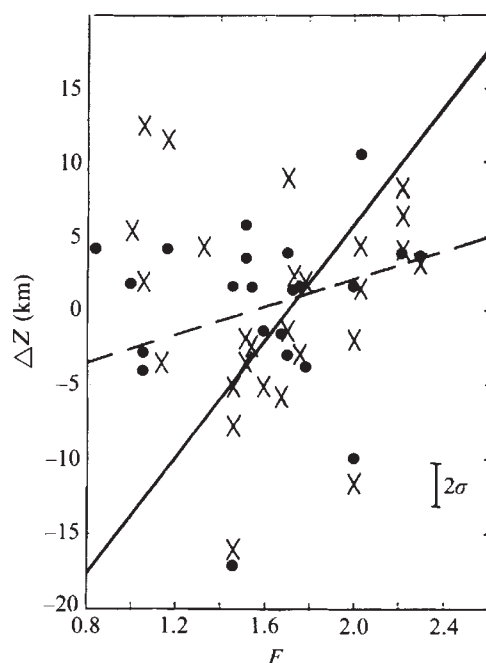


Fig. 6 Residual displacement ΔZ versus non-dimensional tidal range F for western Irish Sea lines 3 and 4 combined. (●, line 3; ×, line 4.) The extent of equilibrium adjustment is shown by the continuous line. The dashed line has the largest slope consistent with the data points at 95% confidence ($F = 1$ at mean neaps).

it from most other fronts and may contribute to anomalous behaviour. We therefore consider a second example.

The western Irish Sea: In this case density is essentially controlled by temperature. A second difference is that the tidal streams are almost parallel to the front so that the tidal advection signal is weak. After subtracting an estimate of ζ based on numerical model results², we have obtained the estimates of ΔZ for both lines 3 and 4 which are plotted against tidal range F in Fig. 6. As in the case of the Islay front, there is no evidence of semi-monthly adjustment. The most probable neaps to springs adjustment is 0.7 km with an upper bound, at 95% confidence, of 4.5 km.

The relatively large scatter is again probably due to frontal meanders and eddies (see Fig. 1b) which frequently occur on this front.

Seasonal and other adjustments

To test for frontal movement on a longer time scale, we have plotted ΔZ versus Julian day number n in Fig. 5c, together with an estimate of adjustment based on the mean nett heat flux through the sea surface. The scatter of points is such that a seasonal trend cannot be reliably identified with the present data, particularly as there are very few data points after $n = 240$ when the seasonal effect may become conspicuous. Moreover, because of the stored potential energy deficit, the retreat of the front may be slower than predicted by an equilibrium model especially as $Q \rightarrow 0$ at the autumnal equinox. Note also that the sampling scheme is biased, in the sense that images are most often available during anticyclonic weather conditions when surface heating will be above average. This may account for the large cluster of points with low values of ΔZ occurring between days 220 and 240. A large proportion of these points are from images during a period of exceptional fine weather in August 1976.

If the observed displacements from the mean position are not semi-monthly or seasonal in origin, to what, apart from frontal instabilities, could they be attributed? One possibility is that they represent movement induced by large scale advection. The residual velocity field in the Irish Sea, averaged, over periods of ~ 1 month, is small and not known with confidence, but over shorter time scales, residual flows of a few centimetres per second probably occur⁷ and these could be responsible for fluctuations in the frontal position of the observed magnitude of $\approx \pm 5$ km.

To assess the relevance of this hypothesis, we have examined the data for coherent adjustment on adjacent reference lines. No significant correlation is found for example, between ΔZ on lines 1 and 2 for the Islay front, implying that other movements with a length scale shorter than the line spacing of 30 km, are dominant.

Discussion

The results reported here indicate that any adjustment of the shelf sea fronts to the semi-monthly mixing cycle is minimal. There is also little support in the data for the notion of seasonal positional adjustment between June and the end of August. The shelf sea fronts therefore represent a set of almost fixed geographical boundaries throughout the summer season, moved significantly only by tidal advection.

There remains the possibility of regular adjustment on a very small spatial scale ~ 2 km, which cannot be resolved by the present approach, but which may be of considerable biological importance. It may be significant that the three regression coefficients of ΔZ versus F , although small, are all positive, which is the sense of the anticipated adjustment.

The remarkable consistency of frontal positions confirms the prediction of the James model and points to the operation of a feedback process in determining the rate of vertical mixing, that is, once stratification is established the efficiency of mixing falls, so that even when stirring increases at spring tides, the stratified structure survives. The notion of feedback is also consistent with the results of mixing tank experiments⁸ which indicate that the flux Richardson number (efficiency) decreases with stability above a threshold stability level. An interesting experiment described by Loder (unpublished) suggests that direct laboratory study of this hysteresis effect may be profitable.

Received 30 April; accepted 11 July 1979.

1. Simpson, J. H., Allen, C. M. & Morris, N. C. G. *J. geophys. Res.* **83**, 4607-4614 (1978).
2. Pingree, R. D. & Griffiths, D. K. *J. geophys. Res.* **83**, 4615-4622 (1978).
3. Simpson, J. H. & Hunter, J. R. *Nature* **250**, 404-406 (1974).
4. Pingree, R. D. *J. mar. biol. Ass.* **58**, 955-963 (1978).
5. James, I. D. *Estuar. coast. mar. Sci.* **5**, 339-353 (1977).
6. Simpson, J. H., Edelsten, D. J., Edwards, A., Morris, N. C. G. & Tett, P. B. *Estuar. coast. mar. Sci.* (in the press).
7. Booth, D. thesis, Univ. Wales (1978).
8. Linden, P. F. *Geophys. astrophys. Fluid Dynamics* (in the press).

The nitrogen budget of a salt marsh ecosystem

Ivan Valiela & John M. Teal

Boston University Marine Program, Marine Biological Laboratory, Woods Hole, Massachusetts, and Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Salt marshes reduce oxidised nitrogenous compounds to ammonium and particulate nitrogen and export these reduced forms to coastal waters. The internal demands exceed the net inputs of nitrogen by rain, groundwater flow and fixation, suggesting very active uptake, conversion, release and recycling of nitrogen within a marsh ecosystem. Nitrogen losses are mainly through tidal exchanges and denitrification, and these two outputs balance the gains. The chemical exchanges among uplands, marshes and coastal water are important in structuring these ecosystems.

NITROGEN is a key nutrient in coastal ecosystems¹. In salt marshes, nitrogen has a critical role in determining the function and structure of the ecosystem. Salt marsh florae receiving increased amounts of nitrogen show increased rates of primary production²⁻⁵, the abundance of certain species changes and the nitrogen content of the plants increases⁶. Provision of excess nitrogen raises the per cent nitrogen in tissues of marsh grasses from levels typical of C_4 plants to those of C_3 plants^{7,8}. This is important because low nitrogen content is one reason why C_4 plants are relatively free from herbivores⁹. The dominant plants (*Spartina alterniflora*, *S. patens* and *Distichlis spicata*) in salt marshes of the Atlantic coast of the US are C_4 grasses, but where the supply of nitrogen is increased the number of marsh herbivores increases due to the lowered C/N ratios in the plant tissues⁶.

Increases in nitrogen supply also lead to increases in the activity and standing crops of decomposer organisms and detritus-feeding invertebrates as the percent nitrogen of the litter helps to regulate decomposers (unpublished results). In general, the biomass of invertebrates in salt marsh plots receiving added nitrogen is larger than in untreated plots. We have also found that increased nitrogen supply changes the morphology of stands of *Spartina alterniflora* from a short to a tall form¹⁰. Concomitant with this is increased spacing among the taller plants leading to a more open habitat which in turn allows predatory fish and crabs a greater chance of finding prey on the marsh surface¹¹. Increases in the supply of nitrogen therefore enhance primary production, decomposer activity, secondary production and changes in the physical structure of a marsh

environment that facilitate predation by fish. Thus the rate of turnover of the entire system to a large extent depends on nitrogen supply.

A second important feature of nitrogen in salt marshes involves the export of nitrogenous materials and marsh-produced organic matter to coastal waters. The proposition that marshes provide food for coastal and estuarine fin and shell fish populations, including economically important species, has been the main argument for conservation of coastal marshes. With the recognition that nitrogen has a limiting role for coastal phytoplankton production¹, it can be appreciated that nitrogen exchange between coastal waters and marshes is significant for both these ecosystems.

We present here the first complete nitrogen budget for a marine ecosystem where each of the various inputs and outputs from a salt marsh has been accounted for. Great Sippewissett Marsh, our study marsh, is located on the western shore of Cape Cod, Massachusetts. Seawater from Buzzards Bay flows into Great Sippewissett Marsh through a single entrance (Sippewissett Creek) and floods the marsh twice daily with a maximum tidal excursion of about 1.6 m. The total area of the marsh is 483,800 m² and includes muddy (104,900 m²) and sandy (61,700 m²) creek bottoms, algal mats (13,000 m²), low marsh dominated by short (122,500 m²) or tall (91,100 m²) *Spartina alterniflora* and high marsh (89,300 m²) where the grasses *S. patens* and *D. spicata* are found. The marsh is surrounded by glacial moraine on the south, east and north and by sand dunes to the west. The relief and the ground water table slope generally to the west and towards the marsh and bay.

Inputs and outputs

Nitrogen enters Great Sippewissett Marsh through precipitation and flow of groundwater, the latter flowing into the marsh through numerous small underground springs. Other sources of nitrogen are fixation of atmospheric nitrogen by blue-green algae and bacteria in the sediments. Tidal flushing brings in and removes dissolved and suspended nitrogen from the marsh. The second major loss of nitrogen is through denitrification. Other minor fluxes include volatilisation of ammonia, removal of shellfish, loss to sediments and deposition of faeces by roosting birds.

Groundwater flow. Groundwater was sampled over 2 yr at seven small underground springs that flowed from the upland into the marsh¹². The amounts of ammonium, nitrite and organic nitrogen were determined using an autoanalyser. The total flow of groundwater into the marsh lowered the salinity of the ebbing tidal water by about 2‰. This lowered salinity was used to calculate the volume of freshwater contributed by groundwater. Salinities were collected routinely during entire tidal cycles through the year. With these estimates of concentrations and volumes, the amount of each nutrient entering the marsh at different times of the year was calculated (Fig. 1). Nitrate was the major inorganic nitrogenous constituent of groundwater (averaging over 50 $\mu\text{mol l}^{-1}$), with substantial amounts of dissolved organic nitrogen (DON) (Table 1). There was a small reduction of nitrate flow during the autumn (Fig. 1). The amount

Cover photograph shows false colour IR aerial view of Great Sippewissett Marsh, Massachusetts. The entire system is drained by Sippewissett Creek. The dark blue tidal creeks have muddy bottoms while sandy bottoms show as light blue. Small pannes with no obvious drainage are scattered through the marsh. Algal mats occur on sand flats (top right). The sand surface is very unstable except where consolidated by the grey-blue mats of blue-green algae and purple sulphur bacteria. The largest area of the marsh is low marsh, covered by *Spartina alterniflora* (dark pink). The taller stands of *S. alterniflora*, showing as a thin red ribbon, can be seen on creek banks. Above the elevation of the low marsh, visible as areas of lighter pink, is high marsh where *S. patens* and *Distichlis spicata* grow. The marsh is surrounded by upland vegetation on the glacial hills.

of ammonium increased in summer, presumably due to the increased decay of litter prompted by higher temperatures at that time of year. The amounts of nitrite involved were minimal and were not included in Fig. 1. The total annual input of nitrogen to Great Sippewissett Marsh by groundwater was 6,120 kg yr⁻¹.

Precipitation. The amount of precipitation and the concentration of ammonium, nitrate, nitrite, DON and particulate nitrogen were measured in rain gauge collections at various sites over 2 yr (ref. 12). Nutrient inputs for the four seasons were calculated by multiplying the concentrations of nitrogenous nutrients by the amount of rain for each rainfall. The nutrient concentrations were typical of semirural areas¹³⁻¹⁵ and there was little seasonal change. Nitrate and DON were the principal forms of nitrogen delivered by rain (Fig. 2 and Table 1). The concentration of all nutrients was higher in the groundwater than in precipitation¹² primarily due to evapotranspiration by terrestrial vegetation. There is little evidence of any increase in nitrogenous concentration in groundwater due to leaching from septic tanks upstream. The volume of rain that fell on the marsh was much less than that which reached the marsh as groundwater, so only 380 kg yr⁻¹ of nitrogen reached the marsh through rainfall (Table 1).

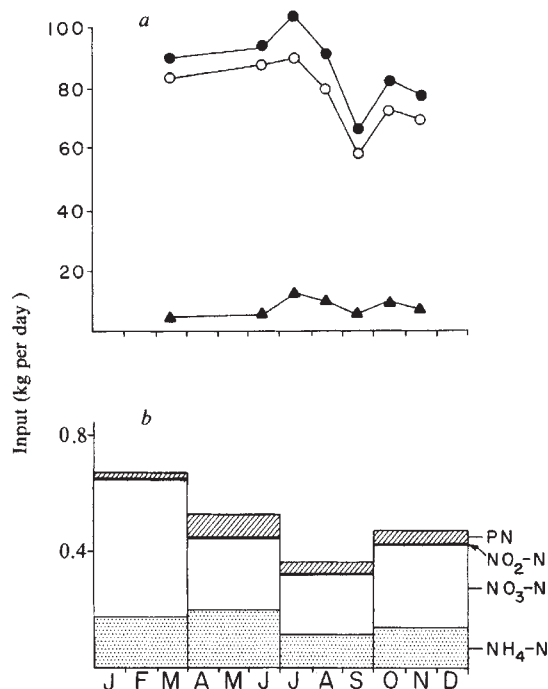


Fig. 1 Amounts of nitrogenous nutrients brought in by groundwater flow (a) and precipitation (b) into Great Sippewissett Marsh throughout the year. Values for precipitation are averaged for 3 months periods. Even though rain fell in episodic fashion, results are averaged on a per day basis over an entire 3-month period for ease of comparison with the other inputs and outputs. ●, Dissolved organic nitrogen; ○, NO₃-N; ▲, NH₄-N.

Nitrogen fixation by bacteria. Fixation of nitrogen by bacteria in the marsh is due to free-living bacteria in sediments lacking vegetation and bacteria closely associated with roots of marsh plants¹⁶. Rates of bacterial fixation in each of the marsh habitats were measured by incubating cores of sediment for 24 h in a gas mixture of 86% nitrogen, 10% acetylene, 4% oxygen. We then refilled the headspace above the core with gas of a similar composition and measured ethylene production 24 h later by gas chromatography¹⁷.

Low marsh was the major site of bacterial fixation due to the higher rates (100–500 ng N cm⁻² h⁻¹) during the warm part of

the year) and larger area involved (Fig. 2). High marsh had lower rates (50–150 ng N cm⁻² h⁻¹) and smaller area. Fixation peaked in the vegetated habitats during spring and early autumn, when roots were actively growing or senescence was taking place. Perhaps root exudates are most abundant at those times. Fixation rates in creek bottoms were considerably lower (maximum of 80 and 20 ng N cm⁻² h⁻¹ for muddy and sandy creeks) than in those recorded in sediments supporting plants and peak fixation occurred in midsummer (Fig. 2). Nitrogen fixation totalled 2,980 kg yr⁻¹.

Table 1 Nitrogen budget for Great Sippewissett Marsh (kg yr⁻¹)

Processes	Input	Output	Net exchange
Precipitation	380		380
NO ₃ -N	110		
NO ₂ -N	0.4		
NH ₄ -N	70		
DON	190		
Particulate N	15		
Groundwater flow	6,120		6,120
NO ₃ -N	2,920		
NO ₂ -N	30		
NH ₄ -N	460		
DON	2,710		
N ₂ fixation	3,280		3,280
Algal	297		
Rhizosphere bacteria	2,595		
Non-rhizosphere bacteria	384		
Tidal water exchange	26,200	31,600	-5,350
NO ₃ -N	390	1,210	
NO ₂ -N	150	170	
NH ₄ -N	2,620	3,540	
DON	16,300	18,500	
Particulate N	6,740	8,200	
Denitrification		4,120 (2,820)	-6,940
Sedimentation		1,295	-1,295
Volatilisation of NH ₃		17	-17
Deposition of bird faeces	9		9
Shellfish harvest		9	-9
Totals	35,990	39,860	-3,870

Totals for each process are offset to the right. Losses from the marsh to Buzzards Bay are shown as negative numbers in the net exchange column. The amount of nitrogen added to the denitrification value is not equal to the amount of nitrogen fixed because net denitrification was not active in the blue green algal mats. Some of the values in this table and in the text may not agree with those of refs. 12, 16, 18, 20. The entries here were computed using corrected areas for the habitats and supersede earlier estimates.

Nitrogen fixation by algae. Algal fixation was measured for each habitat by an adaptation of the acetylene reduction technique¹⁸. Although pannes and algal mats had the highest fixation rates (up to 500–600 ng N cm⁻² h⁻¹), the area of these habitats was so small that their contribution of N to the total marsh was small. Low marsh and creek bottoms showed lower rates¹⁸ but their larger areas resulted in a greater contribution to the nitrogen economy of the marsh. The total nitrogen introduced into the marsh by algal fixation was 297 kg yr⁻¹, an order of magnitude smaller than that contributed by bacterial fixation.

Tidal exchange. The flow of seawater and nitrogenous nutrients through the mouth of Sippewissett Creek were measured during entire tidal cycles about monthly throughout the year. Water movement in and out of the marsh was measured with mechanical flow meters throughout each tidal cycle¹². Tidal water was sampled hourly during each cycle and the concentrations of dissolved and particulate nutrients were determined. With these data and knowledge of the cross-sectional area and tidal height of Sippewissett Creek, water flux and nitrogenous nutrients entering and leaving the marsh during each hour of a tidal cycle were calculated. We summed over the tidal cycle to obtain net exchange (Fig. 3a).

Ammonium was the principal form of dissolved inorganic nitrogen (DIN), averaging about 1 μmol NH₄-N l⁻¹ in winter and spring and 2–6 μmol NH₄-N l⁻¹ in summer and autumn. For the first half of the year there were only trivial exchanges of

ammonium between the marsh and Buzzards Bay (Fig. 3a). From June to July there was a substantial import of ammonium into the marsh. This took place while the vegetation was actively growing and setting seed. The growth rate of *S. alterniflora* in Great Sippewissett Marsh is such that this species by itself consumes more than 39 kg per day of $\text{NH}_4\text{-N}$. During the same time of year about 9 kg per day of NO_3^- and NH_4^+ nitrogen are brought into the marsh daily by groundwater (Fig. 3b). The tidal import of DIN into the marsh (Fig. 3b) was 8 kg per day, making a daily total of 17 kg DIN per day, a rate lower than the 39 kg per day used by plants during their period of active growth.

Towards the end of August there was a marked export of ammonium into Buzzards Bay (Fig. 3a), related to the maturity and senescence of marsh vegetation. The peak export of ammonium from the marsh was after the plants had matured and began flowering (Fig. 4); mature and senescent plants are most subject to leaching¹⁹. In a salt marsh with twice daily submergence in a good electrolyte, leaching should be important. Measurements of leaching of $\text{NH}_4\text{-N}$ from marsh grasses show that such losses peak in August and that the standing crop of marsh grasses can release about 7 kg N per day through leaching (unpublished results).

In late summer the DIN conveyed into the marsh by groundwater (about 8 kg N per day), is no longer intercepted by the marsh ecosystem. Adding this amount of groundwater DIN to the 7 kg N per day leached from plants, a value (15 kg N per day) that is near the observed export of DIN (roughly 12 kg N per day) from the marsh to the Bay in late August and September is obtained (Fig. 3b). Perhaps additional ammonium is released by anaerobic respiration of plant exudates. In August the sediments become markedly more anoxic (B. Howes, personal communication). We are investigating the relationship between root activity and redox condition but as yet are unable to estimate the contribution of this mechanism to tidal export of DIN.

Nitrite is found in low concentration (up to $0.2\text{ }\mu\text{g-at NO}_2\text{-N l}^{-1}$) in marsh water and it is difficult to know if the fluctuations seen in Fig. 3 (top) are of significance. Nitrate concentrations were higher (averaging up to $1\text{ }\mu\text{mol l}^{-1}$), with net exports in autumn and winter (Fig. 3a).

Figure 3b shows the seasonal changes of total dissolved inorganic nitrogen. The marsh intercepts most of the groundwater DIN (mostly nitrate) during winter and spring. This may be due to the conversion of nitrate dissolved in the water into nitrogen gas by denitrifying organisms, described later. During midsummer the marsh imports DIN, as the grasses take up ammonium during maximum growth. With the onset of senescence, plant leaching and decay lead to export of DIN. The coming of winter slows all these processes. Throughout most of the year the marsh intercepts DIN (principally nitrate) that otherwise would flow into deeper waters.

Denitrification. In low and high marsh—habitats where plant roots consolidate sediments—rates of denitrification were measured using a bell jar method. Darkened glass bell jars were pushed about 20 cm into the sediment and the headspace was flushed with helium. Argon was injected into the headspace as an internal standard. After a 4-h incubation, samples of the gas in the headspace were removed and the ratio of nitrogen-to-argon was measured by gas chromatography.

The denitrification rates of the wet flocculent sediment of creek bottoms and pannes were assayed using a gas-partitioning technique. Samples of the sediment were placed in serum vials and flushed with helium. After incubation for 24 h the gas in the sediment was stripped onto the headspace by shaking. The gases were then sampled and analysed by gas chromatography²⁰.

The rates of denitrification were measured monthly at each of the habitats (Fig. 2). Due to the higher rates (maximum of $5\text{ mg N m}^{-2}\text{ h}^{-1}$) and large area of creek bottoms²⁰, this habitat was the site of over 50% of the net denitrification for the marsh. The low marsh had peak denitrification rates of $2\text{--}3\text{ mg N m}^{-2}\text{ h}^{-1}$, while the high marsh rates were only $\sim 1\text{ mg N m}^{-2}\text{ h}^{-1}$. The rates in the pannes reached over

$5\text{ mg N m}^{-2}\text{ h}^{-1}$ but their small surface area did not make a large contribution to losses of N from the marsh as a whole. In the algal mats the high fixation rates were matched by high rates of denitrification²⁰. In all habitats there was a strong seasonal pattern associated with the annual temperature cycle.

Both techniques used to measure denitrification yield results that represent the difference between true denitrification rates and simultaneous nitrogen fixation within the sediment. In Table 1 we added the measured amounts of nitrogen fixed to the

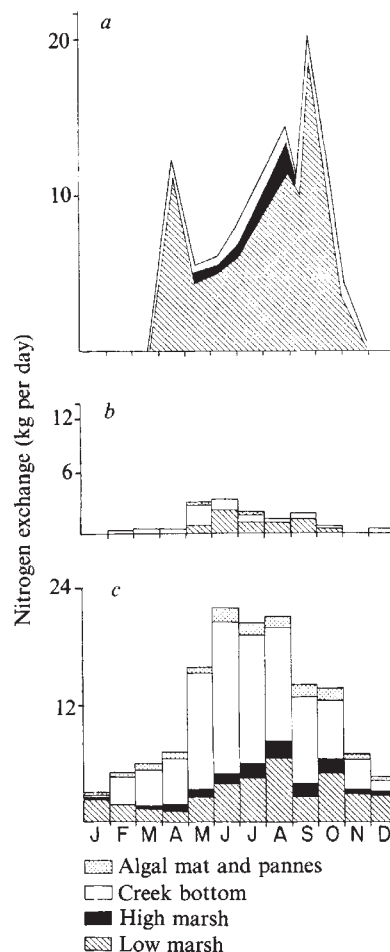


Fig. 2 Nitrogen exchanged by microbial processes in Great Sippewissett Marsh. Each major habitat is shown separately. Values are the products of rates $\times$ areas of each of the habitats. a, Bacterial N_2 fixation; b, algal N_2 fixation; c, denitrification.

amount of nitrogen lost by denitrification to obtain the total amount denitrified. This rate of total denitrification shows that low marsh and creek bottoms are the habitats where denitrifiers are most active and contributed the most to losses of N by denitrification (Fig. 2). The other habitats are of minor importance and the whole marsh exports about $6,940\text{ kg N yr}^{-1}$ as nitrogen gas.

Nitrate is not exported during the spring and summer. This is the time of year when rates of denitrification are highest, suggesting that denitrifiers are active enough to substantially alter the chemistry of tidal water.

Losses to sediments. The rate of accretion of sediments was obtained from measurements made for marshes near Great Sippewissett^{21,22}. For well-established vegetated zones, we used a value of 0.15 cm yr^{-1} as the rate of sediment accumulation. In sediments of Great Sippewissett Marsh there is an average of

1.4% N and the specific gravity of 0.2 g dry weight cm^{-3} . We estimate that in the area stabilised by the growth of vegetation (302,900 m^2), 1295 kg N yr^{-1} is lost to the sediments. Although the loss of nitrogen to sediments is substantial relative to other inputs and outputs (Table 1), the amount of nitrogen sedimented is only about 1% of the estimated 116,800 kg N stored in the top 15 cm of marsh sediments. The bulk of the roots of marsh plants are found above a depth of 15 cm (ref. 23) and therefore any nitrogen found above 15 cm could be potentially recirculated by plant uptake. We are concerned here with losses that occur when nitrogen is left below a depth of 15–20 cm as the marsh peat accumulates.

Minor sources and losses. Our value for the amount of nitrogen lost from Great Sippewissett Marsh through volatilisation of NH_3 is based on rates measured in North Carolina²⁴. At the range of pH values (5–8) measured in our marsh sediments, volatilisation would be small (Table 1).

Flocks of gulls use the marsh as a roost. As they feed largely elsewhere, we consider their defaecation to be an input of nitrogen. We calculated this small input (Table 1) by noting the average number of birds in the marsh, the average defaecations per unit time and the weight and per cent of nitrogen in the faeces.

There is active shellfishing in Great Sippewissett Marsh. The removal of nitrogen by this pathway was estimated by using the records kept by the shellfish warden and the per cent nitrogen in the shellfish meats. This is also a small figure (Table 1).

The balance sheets

Table 1 summarises annual nitrogen transport. The sum of the inputs is remarkably close to the sum of the outputs (only about 11% difference). This balance would be required for long-term persistence of an ecosystem in its present state.

Although adequate replication and estimates of variation for each process are available^{12,16,18,20}, as only one marsh was studied the variability for each entry in Table 1 cannot be easily estimated. There is the further problem that even very small errors in, for example, rates of flow of tidal water may be more important than larger errors in measurements of nutrient concentrations. This is inherent in the arithmetic used in making the estimates of Table 1. From our experience and by pooling estimates of variation from the data available^{12,16,18,20}, we believe that the coefficient of variation of each entry of Table 1 would be about 20% of the mean. Thus, the small 11% variation between inputs and outputs of Table 1 is fortuitous. This agreement suggests that major parts of the nitrogen budget have been accounted for.

Does the marsh have much flexibility to adjust to increased eutrophication of ground and tidal waters? Of the total nitrogen budget of Great Sippewissett Marsh, 91% of the inputs and 83% of the outputs were driven by physical forces which might indicate little room for adjustment. However, if the refractory (see below) dissolved organic nitrogen is not considered, less than 50% of the exchanges are purely physical and the possibility of considerable biological adjustment may be indicated. As dissolved ammonium increases, nitrogen fixation is sharply curtailed²⁵. Fixation amounts to 9 or 20% of the input, again depending on whether or not the DON is considered an inert fraction of the nitrogen budget. In either case this biological feedback system can make only a relatively small change in the system's ability to handle increased nitrogen inputs. Perhaps increased rates of denitrification with increased supply of nitrogen can also serve as a feedback mechanism for adjusting nitrogen flow through the marsh. Our evidence is not conclusive²⁶, although others have shown increased denitrification after nitrate enrichment²⁷. To be effective, increases would have to be large and our rates are already among the higher values reported²⁰. Salt marsh sediments show very large rates of sulphate reduction²⁸. As nitrate is energetically more favourable as an electron acceptor than sulphate, large

increases in nitrate reduction are possible if new supplies of nitrate are made available.

Internal mechanisms, nevertheless, are unlikely to regulate the amount of nitrogen entering salt marshes. Eutrophication of entering fresh and tidal waters will probably result in an increase of the nitrogen pool within the marsh, especially in sediments and vegetation. This will to some extent enhance primary production in salt marshes^{2–5}. Human contamination affects principally ground and tidal water, the two largest processes providing nitrogen to the marsh. Of the two, contamination of groundwater may be more important as flow is one way into the

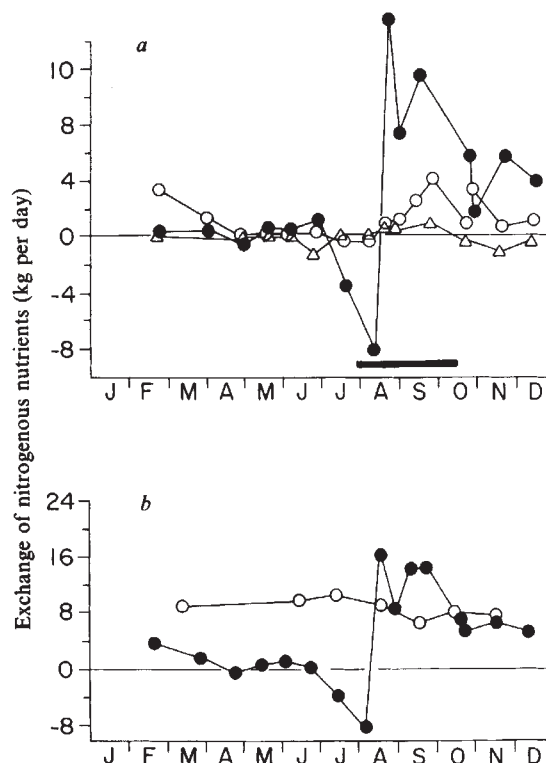


Fig. 3 Exchanges of nitrogenous nutrients between Great Sippewissett Marsh and Buzzards Bay through the year. In *a*, points are net exchanges for a given tidal cycle and recalculated in units of kg per day. ●, $\text{NH}_4\text{-N}$; ○, $\text{NO}_3\text{-N}$; △, $\text{NO}_2\text{-N}$. The solid bar shows the flowering season for *Spartina alterniflora*. *b*, Compares net tidal exchanges of dissolved inorganic nitrogen (●) with the dissolved inorganic nitrogen furnished by groundwater (○) flowing into the marsh. Values in both graphs above the 0 in the y intercept are net exports from Great Sippewissett Marsh, while negative values show an import into the marsh.

marsh and, at least in North America, disposal of sewerage wastes is conducted primarily by leaching wastewater through cesspools and septic tanks into the soil. This may result in enhanced concentrations of nutrients, particularly nitrate, in groundwater throughout the coastal zone.

A better view of the action of a salt marsh ecosystem as a transformer of chemical forms of nitrogen can be obtained by rearranging the entries of the nitrogen budget. The annual exchanges of the major forms of nitrogen for Great Sippewissett Marsh are shown in Table 2. Sixty-four percent of the nitrogen entering the marsh as nitrate is intercepted by the marsh. The oxidised nitrogen imported into the marsh is reduced to various forms within the marsh. Ultimately most of the entering nitrogen is exported by tidal water in the form of particulate, ammonium and molecular nitrogen.

These exports are substantial. Export of dissolved inorganic nitrogen from the area of marshland around Buzzards Bay at a rate comparable to that found in Great Sippewissett Marsh in

Table 2 Annual exchanges of major forms of nitrogen for Great Sippewissett Marsh (kg yr⁻¹)

Form of N	Input	Output	Net	% of input
NO ₃ -N	3,420	1,220	2,200	64
NH ₄ -N	3,150	3,550	-400	-13
DON	19,200	18,500	700	4
Particulate N	6,750	8,200	-1,460	-22
N ₂	3,280	6,940	-3,660	-112
Totals	35,800	38,410	-2,610	-7

The exchanges of nitrite and the nitrogen lost to the sediments are not included.

August could entirely replace the nitrate content of Buzzards Bay in 100 days¹². The amount of particulate organic matter exported yearly from the marsh is equivalent to about 40% of the net annual above ground production of the marsh¹². As salt marshes are among the most productive ecosystems on a per unit area basis, these amounts exported are quantitatively significant and must be important to detritus feeders in coastal waters.

As annual denitrification rates exceed fixation, there are large losses of nitrogen from the marsh as N₂ gas. This is the second major pathway of nitrogen export.

The large amounts of dissolved organic nitrogen (Tables 1, 2) are mostly unaffected by passage through the marsh. This agrees with other results that suggest that the bulk of DON is refractory to microbial attack²⁹, although DON is the principal form of nitrogen in terms of total amounts.

Approximately 6,940 kg of nitrate nitrogen are required to supply the marsh denitrification annually. This is a larger amount than is available through the net import of nitrate (2,200 kg yr⁻¹, Table 2). This means that some nitrification of ammonium must take place within the sediments.

The amount of ammonium entering the marsh is composed of three sources; (1) net exchange by tides, groundwater and rain (1,450 kg N yr⁻¹); (2) the small amount of mineralisation of DON (700 kg N yr⁻¹); and (3) the amount fixed by micro-organisms (3,280 kg N yr⁻¹). The total of these three sources is 5,430 kg N yr⁻¹. After allowing 1,460 kg N yr⁻¹ for export of particulates and 1,295 kg N yr⁻¹ for losses to sediments, only 2,675 kg N yr⁻¹ are available for use within the marsh. Yet we estimate that growth of marsh plants must use about 8,800 kg N yr⁻¹. Thus, it is clear that very active uptake, release, conversion and recycling of nitrogen must take place within the marsh.

If the salt marsh were not present, Buzzards Bay would receive the nitrate from the groundwater and the phytoplankton would be the principal part of the Bay's ecosystem that would be affected. The marsh intercepts much of this nitrate and exports primarily particulate nitrogen to the sea. This export principally affects the animals that feed on particles, a very different component of the coastal food web. These interactions between marsh and bay maintain the structure of benthic and fish communities in the bay. Similarly, the flow of groundwater structures the marsh biota by providing dissolved inorganic nitrogen for plants and denitrifiers.

These results demonstrate that there are significant interactions among units of landscape, in our case nutrient exchanges

among upland, marsh and coastal waters. In view of their importance it is obvious that these interactions among ecosystems have not received enough attention. Many aspects of the structure of an ecosystem, such as what plants, animals or microbes are present, and how many of each kind, are strongly affected by nutrients delivered from outside the particular ecosystem.

Conclusions

Our findings show that there are significant exchanges of nutrients between upland, marsh and coastal waters, and these interactions affect the structure and composition of these systems.

Groundwater often imports more N into salt marshes than does precipitation. Contamination of coastal zone groundwater by sewage is probably important to the marsh nitrogen economy. However, tidal water causes a net export of all nutrients studied, especially ammonium and particulate nitrogen. Quantitatively, tidal exchange is the most important nitrogen transport mechanism. Ammonium export seems to be linked to the activity of marsh grasses, and uptake and leaching are important mechanisms. Nitrogen fixation by bacteria is larger than that by blue-green algae and most bacterial fixation is associated with the roots of marsh grasses.

Dissolved organic nitrogen is the major form of N in the water but little of this is available to decomposers. Despite this rates of denitrification are high. Denitrification may be responsible for the interception by the marsh of the nitrate transported by the flow of groundwater. Annual denitrification for the entire marsh exceeds nitrogen fixed, causing a net export of N₂ to the atmosphere.

Loss of nitrogen to deeper sediments is significant, but this is small in proportion to the nitrogen pool found in the surface sediments where plant roots are found.

The amount of dissolved inorganic nitrogen required to satisfy growth of higher plants and supply the nitrate for denitrification is larger than imports of DIN, which suggests that nitrification, nitrogen fixation and active recycling of N are all important within the marsh. Nitrogen export by the marsh is also important for coastal waters and the transformation of oxidised imports to reduced exports is significant to the coastal ecosystems that receive marsh exports.

Altogether, the imports and exports of nitrogen are in balance. This, along with other evidence, suggests that an old salt marsh such as that studied is in a long-term steady state. Young marshes may be traps for sediment and particulate nutrients suspended in tidal water, but this becomes less marked as the marsh develops, with mature marshes exporting particulates.

We thank S. Volkmann, D. Shafer, C. Van Raalte, W. Kaplan, N. Butler, E. Carpenter, C. Cogswell, D. Berlo, S. Mlodzinska, P. Clarner and P. Glibert for their contributions, Charles Hall, John Hobbie and G. W. Woodwell for reviews of an earlier version of the manuscript. This research was supported by NSF grants GA 43008 and GA 43009. We are indebted to Salt Pond Sanctuaries and to the late Arnold B. Gifford and his family for permission to use their land on Great Sippewissett Marsh.

Received 11 May; accepted 19 July 1979.

- Ryther, J. S. & Dunstan, W. *Science* **171**, 1008-1013 (1971).
- Tyler, G. *Botanisk Not.* **120**, 433-447 (1967).
- Pigott, C. D. *Ecological Aspects of Mineral Nutrition in Plants* (ed. Rorison, I. H.) 25-35 (Butterworth, London, 1969).
- Valiela, I. & Teal, J. M. *Ecology of Halophytes* (eds Reimold, R. J. & Queen, W. H.) 543-563 (Academic, New York, 1974).
- Valiela, I., Teal, J. M. & Persson, N. Y. *Limnol. Oceanogr.* **21**, 245-252 (1976).
- Vince, S. W. thesis, Boston University (1979).
- Wilson, J. R. & Haydock, K. P. *Aust. J. agric. Res.* **22**, 573-587 (1971).
- Black, C. C. *Adv. Ecol. Res.* **7**, 87-114 (1971).
- Caswell, H., Reed, F., Stephenson, S. N. & Werner, P. A. *Am. Nat.* **107**, 465-480 (1973).
- Valiela, I., Teal, J. M. & Deuser, W. G. *Am. Nat.* **112**, 461-470 (1978).
- Vince, S., Valiela, I., Backus, N. & Teal, J. M. *J. exp. Mar. biol. Ecol.* **23**, 255-266 (1976).
- Valiela, I., Teal, J. M., Volkmann, S., Shafer, D. & Carpenter, E. J. *Limnol. Oceanogr.* **23**, 798-812 (1978).
- Eriksson, E. *Tellus* **4**, 215-232 (1952).
- Allen, S. E., Carlisle, A., White, E. J. & Evans, C. C. *J. Ecol.* **56**, 497-504 (1968).
- Gillam, J. W., Daniels, R. B. & Lutz, J. F. *J. env. Geol.* **3**, 147-151 (1974).
- Teal, J. M., Valiela, I. & Berlo, D. *Limnol. Oceanogr.* **24**, 126-132 (1979).
- Patriquin, D. G. *Environmental Role of Nitrogen-Fixing Blue Green Algae and Asymbiotic Bacteria*, (ed. Granhill, U.) (Bull. Ecol. Res. 26, Stockholm, 1977).
- Carpenter, E. J., Van Raalte, C. D. & Valiela, I. *Limnol. Oceanogr.* **23**, 318-327 (1978).
- Tukey, H. B., Jr. *A. Rev. Pl. Physiol.* **21**, 305-324 (1970).
- Kaplan, W., Valiela, I. & Teal, J. M. *Limnol. Oceanogr.* **24**, 726-734 (1979).
- Redfield, A. C. *Ecol. Monogr.* **42**, 201-237 (1972).
- Redfield, A. C. & Rubin, M. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1728-1735 (1962).
- Valiela, I., Teal, J. M. & Persson, N. Y. *Limnol. Oceanogr.* **21**, 245-252 (1976).
- Raps, M. E. A. *Rep. to Office of Sea Grant* (eds Kuenzler, E. J. & Chestnut, A. F.) (University of North Carolina, Chapel Hill, 1971).
- Van Raalte, C. D., Valiela, I., Carpenter, E. J. & Teal, J. M. *Est. Coast. Mar. Sci.* **2**, 301-305 (1974).
- Kaplan, W. thesis Boston University (1977).
- Van Kessel, J. F. *Water Res.* **11**, 259-267 (1977).
- Howarth, R. W. thesis MIT/Woods Hole Oceanographic Institution (1979).
- Abd Aziz, S. A. & Nedwell, D. B. *Ecological Processes in Coastal Environments* (eds Jefferies, R. & Davy, A. J.) (Blackwell, London, in the press).

Phenotype stabilisation and integration of transferred material in chromosome-mediated gene transfer

L. A. Klobutcher

Department of Human Genetics

F. H. Ruddle

Department of Biology, Yale University, New Haven, Connecticut 06520

The human genes for thymidine kinase, galactokinase and procollagen type I, located on chromosome 17, have been transferred to cultured mouse cells. In unstable cell lines the transferred genetic material is shown to exist independently, whereas in stable cell lines the transferred material has become associated with chromosomes of the recipient cell. The size of the transferred genetic material exceeds 1% of the human haploid genome in some of the transformed cell lines. In addition, the data indicate a gene order of: centromere–galactokinase–(thymidine kinase, procollagen type I), on the long arm of human chromosome 17.

PURIFIED metaphase chromosomes have been used as vectors for the transfer of genetic information between cultured mammalian cells^{1–7}. This process, referred to as chromosome-mediated gene transfer (CMGT), typically results in the transfer of subchromosomal amounts of genetic material^{2–4}. Expression of the transferred genetic material (transgenome) can be either stable or unstable following return of gene transfer cell lines (transformants) to non-selective culture conditions^{1,3,5–7} (for detailed reviews of the CMGT process see refs 8, 9).

To examine the nature of the CMGT process, we have transferred the human chromosome 17 gene thymidine kinase (TK, EC 2.7.1.75) to cultured mouse cells. In some instances co-transfer of the syntenic genes galactokinase¹⁰ (GK, EC 2.7.1.6) and type I procollagen (PCI)¹¹ was detected. This three-gene segment has provided a means of defining the length of a transgenome and facilitated examination of the fate of the transgenome. In addition, we present data concerning the order of these three genes and assess the usefulness of CMGT as an intergenic mapping procedure. We have also observed donor chromosome fragments and co-transfer of asyntenic genes in some transformants.

Generation and characterisation of primary transformants

Human chromosomes isolated from HeLa S3 (TK⁺) cells were applied to the mouse cell line LM(TK[−]) which contains a non-reverting (<10^{−9}) mutant allele of TK. The CMGT procedure used³ produced transformed colonies at an overall frequency of 2.3 × 10^{−7} per recipient cell (Table 1). In control experiments without donor chromosomes no colonies were observed.

Seven independent cell lines, designated TGT, were isolated from two transfer experiments. Two of the transformants,

2TGT2 and 2TGT4, were found to express the human form of GK (Table 2). Human PCI production was assayed by Ouchterlony double immunodiffusion using an antibody to the purified human protein¹¹. Of the seven cell lines examined, only 2TGT4 expressed human PCI.

The stability of the transferred phenotype was assayed by cultivating the transformants in non-selective medium and periodically determining the proportion of cells able to form colonies in hypoxanthine, aminopterin and thymidine (HAT) selective medium (see Fig. 2 legend). Transformants TGT1, 2TGT1, 2TGT2 and TGT3 expressed the TK⁺ phenotype stably, that is, >90% of the cells remained HAT resistant after 30 d in non-selective medium. Transformants 2TGT4 and 2TGT5 were unstable, losing the HAT-resistant phenotype with first order kinetics at a rate of 2.5% per cell generation. The remaining transformant, TGT2, seemed to be a mixture of stable and unstable cells.



Fig. 1 Alkaline Giemsa¹² stained metaphase spread of transformant 2TGT4 containing an independent human chromosome fragment (arrow). Human chromosomal material appears lightly stained, whereas mouse chromosomes stain darkly, with the exception of their centromeric regions (lightly stained). Twenty metaphase spreads were observed and photographed for each cell line analysed.

Karyotypic analyses of the transformants were carried out using the alkaline Giemsa staining technique¹². This method produces a differential staining of rodent and human chromosomes and is effective in detecting small amounts of human chromosomal material in a predominantly mouse background. Of the seven transformed cell lines, two possessed detectable human chromosome fragments. Transformant TGT2 contained a large human fragment that was translocated to a

Table 1 Summary of TK transfer experiments

Expt	No. of LM(TK ⁻) recipient cells	Cell equiv. of HeLa chromosomes	No. of flasks with colonies	Transfer frequency
1	2×10^7	3×10^7	2/20	1×10^{-7}
2	1×10^7	1×10^7	5/10	5×10^{-7}
3	2×10^7	0×10^7	0/20	0

Mitotic HeLa cells were obtained by treating cultures with 5 atm. nitrous oxide for 18 h. Chromosomes were isolated and applied to LM(TK⁻) cells according to the method of Willecke and Ruddle³. Following incubation, recipient cells were diluted with minimal essential medium (α MEM) plus 10% fetal calf serum, and 1×10^6 cells were plated per 75-cm² Corning culture flask. HAT selection was begun 24 h later²⁴. Transformed colonies appeared after 2–3 weeks and were expanded for further analysis. All cell lines were negative for mycoplasma contamination by the culture method of Hayflick as modified by Barile *et al.*²⁵. For the calculation of transfer frequency, each flask containing colonies was considered as one event, irrespective of the number of colonies present.

mouse chromosome. Transformant 2TGT4 contained an independent human fragment (Fig. 1). This fragment was similar in size to the long arm of human chromosome 17 and possessed a terminally constricted centromere-like region. G-banded chromosome preparations¹³ revealed that this chromosome fragment possessed the banding pattern characteristic of the long arm of human chromosome 17. These observations of cytologically detectable chromosome fragments confirm previous work from our laboratory¹⁴. It now seems that there is a range of transgenome sizes, the largest exceeding 1% of the human haploid genome.

Co-transfer of asyntenic genes

The transformed cell lines were analysed for the expression of 20 additional human isozymes assigned to 16 different human chromosomes^{15,16}. Five of the seven transformants were negative for all human isozymes examined. One cell line, TGT2, expressed the human form of adenylate kinase (AK1, EC 2.7.4.3), which has been assigned to human chromosome 9, and another transformant, 2TGT1, expressed human mannose phosphate isomerase (MPI, EC 5.3.1.8), a human chromosome 15 marker enzyme (Table 2).

Two types of experiments indicate that the asyntenic genes have become associated with the selected TK gene. First, nine independent subclones of TGT2 were generated under selective pressure for the TK gene. All these subclones continued to express the human form of AK1. Second, transformed cell lines which had co-transferred both syntenic and asyntenic genes were back-selected against TK by growth in medium with 5-bromodeoxyuridine (Table 2). Mass populations of surviving cells were analysed to reduce the effect of spurious dual segregation events. Back-selectants of transformants 2TGT2 and 2TGT4 lost the linked human chromosome 17 genes concomitant with TK. Similarly, back-selectants of 2TGT1 and TGT2 lost the natively asyntenic genes MPI and AK1. Karyotypic analysis of back-selectants derived from lines containing human chromosome fragments demonstrated loss of the fragments coincident with loss of the transferred markers.

There are at least two mechanisms which could result in the transfer and association of asyntenic genes with selected genes. The HeLa cell line used as the chromosome donor is typical of continuous tissue culture lines in possessing some chromosome rearrangements. Rearrangements involving chromosomes 17, 9 and 15 could have created artificial syntenic groups in the donor cell line which were subsequently picked up in the transfer experiments. Alternatively, the CMGT process could cause co-transfer of genuinely asyntenic genes. Transfer of unlinked markers has been observed in bacterial transformation experi-

ments (congression)¹⁷. Such results are due to the cell's ability to ingest and express more than one piece of DNA. Mammalian cells are capable of ingesting more than one chromosome in CMGT conditions¹⁸. Uptake of more than one chromosome followed by partial degradation and reconstruction could explain co-transfer of asyntenic genes.

In other CMGT experiments we have observed five additional transformants expressing normally asyntenic genes (our unpublished results with Miller and Church). The first mechanism proposed would require extensive rearrangement of the HeLa karyotype to explain these results. As the HeLa cell line is basically triploid with a minimal number of rearrangements¹⁹, we favour the second hypothesis. A rigorous test of the hypothesis will require transfer experiments using a primary cell line with a normal diploid karyotype as the chromosome donor. Such experiments are in progress.

The unstable transgenome and its conversion to stability

Various experiments have indicated that the transgenome in stable transformants has become associated with chromosomes

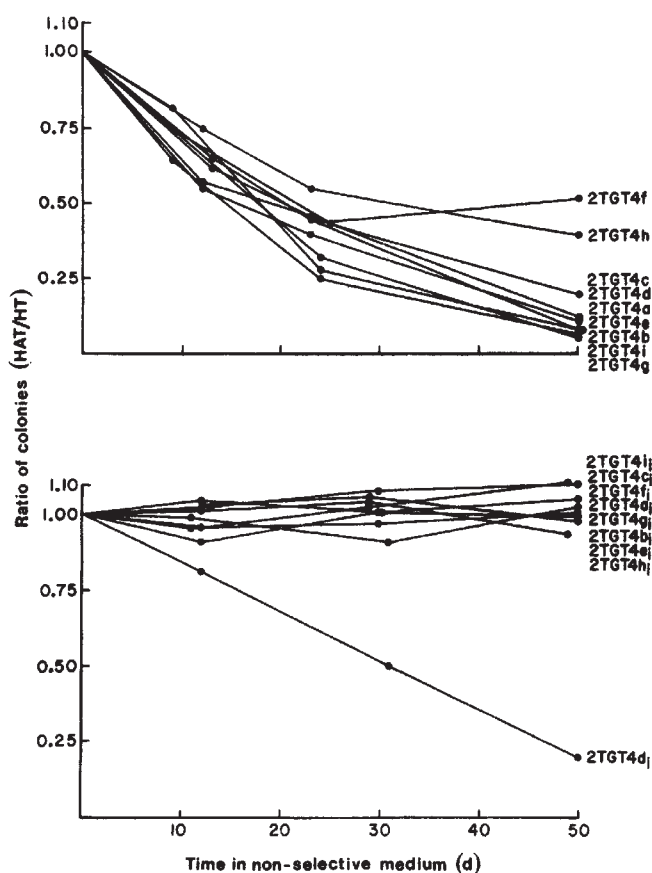


Fig. 2 Stability analyses of subclones of the 2TGT4 transformed cell line. *a*, Stability analyses of nine primary subclones, 2TGT4a–2TGT4i, of transformant 2TGT4. *b*, Second stability analyses of the nine subclones following their growth in non-selective medium for 75 d, and an additional period of growth with HAT-selective medium for 14 d. Eight of the nine subclones now stably express the transferred phenotype. Stability analyses were carried out by culturing transformed cell lines in non-selective medium for at least 30 d. At various times, 200 cells were plated in duplicate flasks with either HAT-selective medium or non-selective HT medium (medium supplemented with 0.1 mM hypoxanthine and 0.016 mM thymidine). After 9–10 d the resulting colonies were stained with 0.25% Wright's stain in methanol and counted. The number of colonies formed in HAT as against HT medium is taken as a measure of the fraction of cells retaining the transferred (TK⁺) phenotype.

of the recipient cell²⁰⁻²³. The precise nature of this association remains unclear. Even less is known about the state of the transgenome in unstable transformants. Prolonged culture in selective medium, however, often results in conversion from unstable to stable expression^{5,7}. It has been postulated that stabilisation occurs infrequently as a unique event in a single cell. The resulting stable cell would thus acquire a growth advantage while in selective medium due to its ability to distribute the transgenome efficiently to daughter cells. Stable cells would then gradually overgrow the unstable cells, giving rise to a completely stable population.

Transformant 2TGT4, which expresses all three human chromosome 17 genes and contains a detectable human chromosome fragment, was chosen for investigating the nature of the unstable transgenome and its conversion to stability. Nine independent subclones (2TGT4a to 2TGT4i) of transformant 2TGT4 were derived and analysed for stability (Fig. 2). Six of the subclones were completely unstable, as more than 90% of the cells lost the HAT-resistance phenotype after 50 d of growth on non-selective medium. Subclones 2TGT4c, 2TGT4f and 2TGT4h did not meet this criterion for unstable expression and will be discussed separately. Each of the six unstable subclones continued to express human GK and PC I. Karyotypic analyses of each of these cell lines revealed a single human chromosome fragment indistinguishable from the fragment present in the 2TGT4 cell line. Thus, the unstable transgenome is an independent and constant entity, losing no genetic information and undergoing no gross structural changes.

Table 2 Summary of TK transformed cell lines

Cell line	TK	GK	PC I	AK1	MPI	Stability	Detectable human chromosomal material
Primary transformants:							
TGT1	+	-	-	-	-	S	-
TGT2	+	-	-	+	-	Mixed	+
2TGT1	+	-	-	-	+	S	-
2TGT2	+	+	-	-	-	S	-
2TGT3	+	-	-	-	-	S	-
2TGT4	+	+	+	-	-	U	+
2TGT5	+	-	-	-	-	U	-
Back-selected transformants:							
TGT2BS	-	ND	ND	-	-	ND	-
2TGT1BS	-	ND	ND	-	-	ND	ND
2TGT2BS	-	-	ND	ND	ND	ND	ND
2TGT4BS	-	-	-	ND	ND	ND	-

Human PC I expression was assayed by Ouchterlony double immunodiffusion using an antibody raised against purified human type I procollagen¹¹. Human GK activity was detected by the starch gel electrophoretic method of Nichols *et al.*²⁶. In addition to human AK1 and MPI, all transformants were assayed for the expression of 18 other human isozymes¹⁵. Back selection of primary transformed cell lines was carried out by growing cells in 30 $\mu\text{g ml}^{-1}$ of 5-bromodeoxyuridine for 24 h. Culture medium was then removed and replaced with phosphate-buffered saline and the cultures exposed to long-wave UV light for 30 min. The cells were returned to growth in culture medium with 5-bromodeoxyuridine, and mass populations of surviving cells were analysed as back-selectants. +, Expression of human phenotype; -, expression of human phenotype not detected; U, unstable expression of transferred phenotype; S, stable expression of transferred phenotype; ND, not determined.

To study the conversion to stability, we isolated stable derivatives of the 2TGT4 subclones. Each of the nine subclones was grown on non-selective medium for 75 d to remove most unstable HAT-resistant cells from the population, and then returned to HAT-selective medium for 14 d to select and fix any pre-existing, small subpopulation of stable cells. The cell populations carried through this regime (designated 2TGT4a_i to 2TGT4i_i) were again analysed for stability (Fig. 2). All the cell

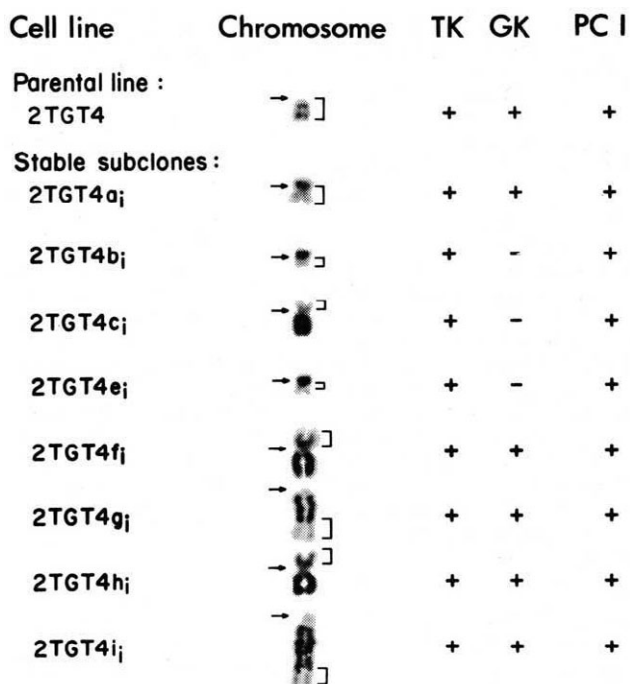


Fig. 3 Expression of human chromosome 17 genes and fate of the human chromosome fragment in the stabilised subclones of transformant 2TGT4. In each case the lightly stained human fragment, indicated by the bracket, has become associated with a darkly stained mouse chromosome. The human chromosome fragment has apparently been translocated to either a mouse chromosome's centromeric region (2TGT4a_i, 2TGT4b_i, 2TGT4c_i, 2TGT4e_i, 2TGT4f_i and 2TGT4h_i) or to the distal portion of a mouse chromosome (2TGT4g_i and 2TGT4i_i). Centromeric regions are identified by the arrows.

lines, with the exception of 2TGT4d_i, then expressed the transferred phenotype stably. Although all eight stable subclones continued to express human PC I, three of the cell lines no longer expressed human GK (Fig. 3). Karyotypic examination revealed that the human chromosome fragment had become associated with a mouse chromosome in all the stable subclones (Fig. 3). In each subclone the fragment was associated with a single mouse chromosome, but the mouse chromosome involved differed between cell lines. The rearrangements appeared either as centric fusions or terminal additions. When metaphase spreads were stained with the fluorescent dye, Hoechst 33258, each of the centric fusion type chromosomes had the brightly fluorescent centromeric region characteristic of mouse centromeric heterochromatin¹³. Neither the independent human fragment in transformant 2TGT4 nor any of its unstable subclones displayed centromeric fluorescence. These observations indicate that the mouse centromeric region has been retained during stabilisation.

Finally, if stabilisation originates in a single cell which then gradually overgrows the population, it should be possible to find cell populations which transiently contain both stable and unstable cells. During a stability test a mixed population would be expected to display a plateau at the frequency of stable cells in the population. The cell lines 2TGT4c, 2TGT4f and 2TGT4h seem to be mixed populations, as they displayed plateaus at 16%, 40% and 50% HAT-resistant cells, respectively (Fig. 2). Karyotypic analysis of these subclones supported this hypothesis. Two types of human chromosome fragment were observed in each of these cell lines. One subset of cells contained the independent human fragment, and a second subset contained the human fragment translocated to a mouse chromosome. The frequency of cells containing the translocated chromosome was similar to the stability test plateau frequency for each cell line. Translocation chromosomes were found in

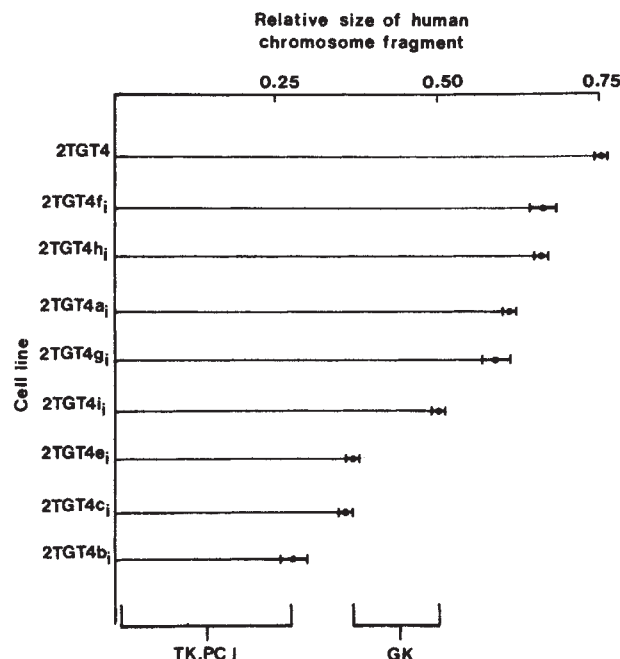


Fig. 4 Histogram displaying the sizes of the human fragments present in the transformed cell line 2TGT4 and its eight stabilised subclones. Relative size values were determined by measuring both the human chromosome fragment and the long arm of a mouse marker chromosome in each metaphase spread. By taking the ratio of the size of the human fragment to the size of the marker chromosome a relative size value was obtained which corrects for variable condensation between metaphase spreads. An average size value was obtained from measurement of at least five metaphase spreads for each cell line. The bars represent ± 1 s.e.m. The regions containing the genes for TK, GK and PC I are indicated at the base of the histogram (see text).

10%, 25% and 35% of the cells of transformants 2TGT4c, 2TGT4f and 2TGT4h, respectively.

These experiments provide a visual record of the stabilisation process. The association of the transgenome with recipient cell chromosomes may be viewed as the acquisition of host centromeric activity, ensuring the orderly distribution of transferred genes at mitosis. Stabilisation does not involve the homologous genetic site in the recipient, as the transgenome is shown to associate with several recipient cell chromosomes. These results agree with those of previous studies^{22,23}. The earlier studies, however, did not rule out the possibility of multiple sites of association within a single cell. We have shown here that each stable cell contains only one copy of the transgenome. Moreover, the fact that each stabilised subclone has only one type of chromosome associated with the transgenome, as well as the observation of mixed populations with regard to stability, argues that stabilisation originates in a single cell which then overgrows the population.

Received 20 February; accepted 2 July 1979.

- McBride, O. W. & Ozer, H. L. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1258–1262 (1973).
- Willecke, G. J., van der Horst, J. & Bootsma, D. *Somatic Cell Genet.* **1**, 137–152 (1975).
- Willecke, K. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1792–1796 (1975).
- Burch, J. W. & McBride, O. W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1797–1801 (1975).
- Willecke, K., Lange, R., Krüger, A. & Reber, T. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1274–1278 (1976).
- McBride, O. W., Burch, J. W. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 914–918 (1978).
- Degen, G. E., Miller, I. L., Adelberg, E. A. & Eisenstadt, J. M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2838–2842 (1976).
- Willecke, K. *Theor. appl. Genet.* **52**, 97–104 (1978).
- McBride, O. W. & Athwal, R. S. *In vitro* **12**, 777–786 (1976).
- Elsevier, S. M. *et al. Nature* **251**, 633–636 (1974).
- Sundar Raj, C. V., Church, R. L., Klobutcher, L. A. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4444–4448 (1977).
- Friend, K. K., Dorman, B. P., Kucherlapati, R. S. & Ruddle, F. H. *Expl Cell Res.* **99**, 31–36 (1976).

Genetic mapping by CMGT

Following stabilisation, three of the 2TGT4 subclones did not express human GK (Fig. 3). Moreover, the length of the human fragment decreased following translocation and was correlated with the number of human genes retained (Fig. 4): the largest fragments were in cell lines retaining all three human chromosome 17 genes, and the smallest were in cell lines retaining only TK and PC I. To examine the nature of the breakage during stabilisation, chromosome preparations were examined by G-banding¹³. All stable transformants, except 2TGT4g_i and 2TGT4i_i, which could not be examined by G-banding due to the similarity of the translocation chromosome to other mouse chromosomes, retained the darkly staining terminal band (17q22–17q25) and various amounts of the remainder of the long arm of human chromosome 17. These results suggest that breaks have occurred near the centromere-like end of the fragment during stabilisation. If we assume that stabilisation involves a single break and fusion of the fragment, and not a complex rearrangement, a deletion map is generated. The map indicates a gene order of centromere–GK–(TK, PC I) (Fig. 4). The data do not allow unambiguous determination of the order of TK and PC I, but failure to segregate PC I suggests that it is distal to TK.

Deletion mapping following stabilisation of visible chromosome fragments provides a means of deriving qualitative intergenic mapping data. Alternatively, the co-transfer frequency of non-selected genes in primary transformants may be a source of quantitative intergenic mapping data^{5,6,9,14}. Although the number of cell lines analysed here is too small to assess directly the usefulness of CMGT in this regard, two observations indicate that only primary unstable transformants should be used to determine a reproducible co-transfer frequency. First, stabilisation coincides with a decreased transgenome size. The co-transfer frequency of a given gene in a population of stable transformants will therefore be lower than in a population of unstable transformants. Second, the consistent loss of the centromere-like end of the chromosome fragment in the stable subclones of transformant 2TGT4 suggests transgenomes may undergo stabilisation in a polar fashion. Polar effects would differentially influence the co-transfer frequency of genes proximal and distal to the selected gene.

Additional transfer lines will have to be examined to determine how generally applicable these findings are. The ability to generate large numbers of transformed cell lines through recent modifications of the transfer technique¹⁴ will aid in both confirming these results and providing quantitative mapping data. The further development of CMGT as a mapping procedure will provide an intermediate level of mapping resolution between restriction endonuclease analysis and somatic cell hybridisation.

We thank Dr R. Church for carrying out procollagen assays, Ms E. Nichols, Mr D. Laska and Ms S. Pafka for technical assistance, Mrs M. Reger and Mrs M. Siniscalchi for preparation of the manuscript, and Drs R. Fournier, G. Scangos, P. D'Eustachio and D. Plotkin and Ms D. Slate for helpful discussions and critical reading of the manuscript. This work was supported by grant GM09966 from the NIH to F.H.R.

- Kozak, C. A., Lawrence, J. B. & Ruddle, F. H. *Expl Cell Res.* **105**, 109–117 (1977).
- Miller, C. L. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3346–3350 (1978).
- Nichols, E. A. & Ruddle, F. H. *J. Histochem. Cytochem.* **21**, 1066–1081 (1973).
- McKusick, V. A. & Ruddle, F. H. *Science* **196**, 390–405 (1977).
- Goodgal, S. H. *J. gen. Physiol.* **45**, 205–228 (1961).
- Simmons, T., Lipman, M. & Hodge, L. D. *Somatic Cell Genet.* **4**, 55–76 (1978).
- Miller, O. J., Miller, D. A., Allderice, P. W., Dev, V. G. & Grewal, M. S. *Cytogenetics* **10**, 338–346 (1971).
- Athwal, R. S. & McBride, W. O. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2943–2947 (1977).
- Willecke, K., Mierau, R., Krüger, A. & Lange, R. *Cytogenet. Cell Genet.* **16**, 405–408 (1976).
- Willecke, K., Mierau, R., Krüger, A. & Lange, R. *Molec. gen. Genet.* **161**, 49–57 (1978).
- Fournier, R. E. K. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3937–3941 (1977).
- Littlefield, J. W. *Science* **145**, 709–710 (1964).
- Barile, M. F., Bodey, G. P., Snyder, J., Riggs, B. D. & Grabowski, M. W. *J. natn. Cancer Inst.* **36**, 155–159 (1966).
- Nichols, E. A., Elsevier, S. M. & Ruddle, F. H. *Cytogenet. Cell Genet.* **13**, 275–278 (1974).

letters

IUE observations of an active region of HD206860

THE chromospheric Mg II *h* and *k* lines seem to have doubly reversed emission core in the spectra of the Sun and of solar and late type stars. In the solar spectrum the general appearance of Mg II *h* and *k* is very similar to the *H* and *K* of Ca II, but with larger intensity fluctuations. Current ideas suggest that the *h* and *k* lines are strongly sensitive to chromospheric parameters. In quiet areas of the Sun the Mg II emission shows asymmetric doubly reversed profile with a brighter blue peak, while in plagues the emission is much stronger and nearly symmetrical¹. The Mg II doublet seems to be a very useful way of investigating the presence of active regions on stars. We report here the results of our observations of the intensity of the Mg II *h* and *k* lines of the star HD206860 collected with the IUE at the Villafraanca Satellite Tracking Station of the European Space Agency (ESA). The aim of the observations was to search for short-term variability of the chromospheric emission connected with the rotation period of the star.

HD206860 is a GOV star with prominent emission cores on the *H* and *K* lines of ionised calcium. Spectrophotometric observations by Wilson² show small amplitude but definite variations of the Ca II *H* and *K* emission components indicating activity in the chromosphere of HD206860. No clear long-term periodicity results from Wilson's observations which were carried out between 1966 and 1975. The data show an apparently random scatter much larger than the scatter of the standard stars. Such scatter may be due to fluctuations of the line intensity in a short time scale.

Since the stars with *H* and *K* chromospheric activity seem the most likely to show photospheric activity, HD206860 was included in a wide programme of UVB photoelectric observations of solar type and late-type stars, undertaken at the Catania Astrophysical Observatory to investigate light variations which might be ascribed to spots or more generally to active photospheric regions^{3,4}. The analysis of the observations from 1970 to 1977 shows regular small amplitude variations in the light of HD206860. The amplitude in V light is 0.02 m and

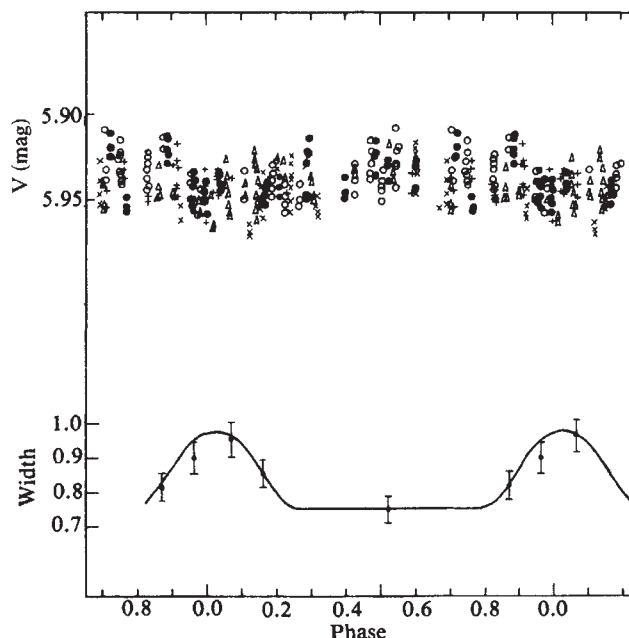


Fig. 2 Photometric observations of HD206860 in V light (different symbols refer to different years of observations) and equivalent width of the *k* emission component of the Mg II versus phases computed with the elements $(JD)_0 = 2440821.48$ and period of 24.9 d.

the period 24.9 d^{4,5}. By the hypothesis that the light variations are due to spots on the star surface, the period of 24.9 d must be the rotation period of the star and this is an appropriate value for a GOV star like HD206860^{5,6}. However, no correlation of the *H* and *K* line data from Wilson's observations with the 24.9 d period was found.

It is surprising that the Ca II *H* and *K* lines do not correlate with the 24.9 d, even if they show short term variability as deduced from the large random scatter in Wilson's observations. Different intrinsic and observational causes may contribute to the lack of the correlation: (1) the variations of the Ca II lines are less prominent than those of the Mg II ones; (2) Wilson's measurements² include photospheric contribution in the line which partially hides the variations of chromospheric origin.

To check these results repeated UV observations of HD206860 were made with the IUE. Five high-resolution spectra ($\sim 0.2 \text{ \AA}$) in the long wavelength region (λ 1,900–3,200 $\text{\AA}$) and two low resolution spectra ($\sim 7 \text{ \AA}$) in the short wavelength region (λ 1,150–1,900 $\text{\AA}$) were obtained from 18 November to 29 December 1978 at the Villafraanca Observatory.

In the long wavelength region we concentrated our attention on the emission cores of the Mg II *h* (λ 2,803 $\text{\AA}$) and *k* (λ 2,796 $\text{\AA}$) resonance lines. A profile of the *k* line in IUE flux units is shown in Fig. 1, where the two emission peaks k_{2v} (the blueshifted component) and k_{2r} (the redshifted component) are evident.

Exposure times (Table 1) have been chosen to produce a well exposed spectrum at the k_2 peak, leading to a signal-to-noise

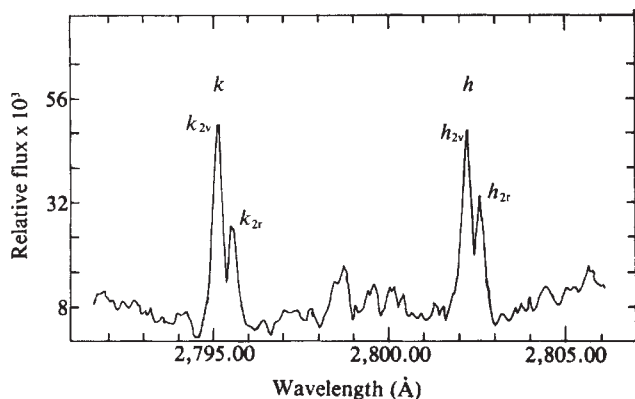


Fig. 1 *h* and *k* Mg II line profiles of HD206860 in IUE flux units. The blue emission component k_{2v} and the red emission component k_{2r} are outstanding.

Table 1 Observational data of HD206860 and measured quantities of the Mg II line referred to the line wing at λ 2,784.5 and λ 2,815

Date (1978)	Exposure time (min)	k_{2V}	k_{2V}/k_{2r}	$(W_k)_E$	ϕ
18 November	44	1.52	1.77	0.814	0.87
23 November	52	1.60	1.55	0.957	0.07
15 December	52	1.74	1.87	0.899	0.96
20 December	48	1.37	1.41	0.854	0.16
29 December	48	1.41	1.66	0.750	0.52

Equivalent widths of the emission $(W_k)_E$ are in Ångströms. The phases ϕ are computed with the elements $(JD)_0 = 2440821.48$ and period of 24.9 d.

ratio of about 20. Three or four pixels were saturated at the k_2 peak on a few spectra. However, this has little influence on the evaluation of fluxes of the whole Mg II emission.

The output of the IUE software, that is the relative intensity spectrum uncorrected, for the wavelength dependence of the scientific instrument sensitivity has been used to study the change of the k lines of Mg II. To reduce the errors, mainly due to the background subtraction, relevant quantities of the emission profile have been referred to two points of the absorption wings at λ 2,784.5 Å and λ 2,815 Å. Since the line wings can be subject to changes due to the activity a check with the continuum at λ 2,685 Å and λ 2,950 Å was made. This check gave no evidence of such changes of the line wings of HD206860. Intensities at k_{2V} and k_{2r} emission peaks and the equivalent width have been measured on the IUE software plots.

Both the k_2 violet and red peaks were found to vary. The largest changes up to 25% were observed for the violet peak. Also the integrated emission component, that is the equivalent width, shows changes of 25%. Equivalent widths of the k line, measured in Angstrom and referred to the average value of the line wing at λ 2,784.5 Å and λ 2,815 Å, are given in Table 1. The estimated error on the equivalent width is $\pm 5\%$. A plot of these values using the period and epoch of the photometric variation, given in Fig. 2, shows a fairly defined variation of the line intensity. Although we have only little data, the increase of intensity near the zero phase seems well defined.

A tentative interpretation of these line intensity changes may be given assuming that they result from a localised bright chromospheric active region like a plage. The continuous line in Fig. 2 represents the change of the k line intensity with a complete rotation of the star for an equatorial bright plage covering 10% of the stellar disk and 2.5 times brighter than the quiet chromosphere. A limb darkening coefficient equal to 0.5 was also assumed. The agreement with the observations seems fairly good.

Solar chromospheric faculae show an enhancement of the k (Mg II) line intensity¹ from two to three times the value of the quiet chromosphere, so the value 2.5 assumed in the case of the active region of the solar type star HD206860 seems to be reasonable.

No accurate value of the limb darkening for the solar k Mg II line is given in the literature but the value 0.12 derived from rocket observations⁷ seems too small to fit the observed behaviour for the variations of HD206860.

The maximum intensity of the Mg II line seems to occur near the minimum of the V light (Fig. 2). The low accuracy of the photometric period and large time interval involved means that this could be only by chance. However, if this correspondence is real, HD206860 is a good example of a star with solar type activity, that is a photospheric dark spot underlying a bright chromospheric plage. So the variations of the intensity of the k line seem to confirm the rotation period of 24.9 d, first determined from the photoelectric observations.

Light variations attributed to spots are known for binary stars of the RS CVn or BY Dra class^{8,9}. From radial velocity observations^{10,11} HD206860 does not appear to be binary, it would therefore be the first single star with an active region detected

from photometric observations and confirmed by the spectroscopic ones.

The two short-wavelength region spectra show spectral features indicating plasma temperatures hotter than 20,000 K. Important emission lines and temperatures of formation are: C II λ 1,335 (20,000 K), Si IV λ 1,394, 1,403 (80,000 K), He II λ 1,640 (80,000 K), C IV λ 1,549 (10^5 K). No clear changes are evident for these lines in the two spectra. Unfortunately, they were obtained at about the same phase on the Mg II variation, that is, during the increasing branch. It would be very interesting to see if any of these high temperature lines disappears or changes its intensity when the active region is not visible.

We thank all the staff of VILSPA Observatory for their assistance in observing and reducing the data and Dr F. Ciatti from Asiago Observatory for the spectrum of 29 December. This work was supported by a CNR-SAS contract.

C. BLANCO
S. CATALANO
E. MARILLI

Osservatorio Astrofisico di Catania,
Istituto di Astronomia dell'Università di Catania,
I-95125 Catania, Italy

Received 9 April; accepted 13 June 1979.

1. Lemaire, P. & Skumanich, A. *Astr. Astrophys.* **22**, 61–68 (1973).
2. Wilson, O. C. *IAU Symp.* No. 71, 447–452 (1976).
3. Blanco, C. & Catalano, S. *Publ. Astr. Soc. Pacific* **82**, 1293–1304 (1970).
4. Blanco, C., Catalano, S. & Godoli, G. *Mem. Soc. Astr. Ital.* **43**, 663–667 (1972).
5. Blanco, C., Catalano, S. & Marilli, E. *IAU. Circ.* No. 1387 (1978).
6. Blanco, C., Catalano, S. & Marilli, E. *Astr. Astrophys.* **36**, 297–306 (1979).
7. Kohl, J. L. & Parkinson, W. H. *Astrophys. J.* **205**, 599–611 (1976).
8. Hall, D. S. *IAU. Coll.* No. 29, 287–318 (1976).
9. Chugainov, P. F. *Izv. Krgm. astrofiz. Obs.* **55**, 94–99 (1976).
10. Shajn, J. & Albitzky, V. *Mon. Not. R. astr. Soc.* **92**, 771–779 (1932).
11. Wilson, R. E. & Joy, A. H. *Astrophys. J.* **111**, 221–261 (1950).

Variations of the Sun's radius and temperature due to magnetic buoyancy

LIVINGSTON¹ has recently measured a decrease in the surface temperature of the Sun coincident with increased solar activity. He interpreted the temperature drop as implying a corresponding reduction in luminosity. I point out here that surface cooling could also be due to a radial expansion of the Sun, with no attendant reduction in luminosity. There is a plausible physical mechanism for such an expansion; namely, variations in magnetic buoyancy due to variations in the magnetic flux in the convection zone over the solar cycle.

The solar luminosity L may be expressed in terms of an effective surface temperature T as $L = 4\pi R^2 \sigma T^4$, where R is the solar radius. For constant R , a decrease in T implies a decrease in L . Alternatively, there can be a cooling (or heating) of the surface due to a radial expansion (or contraction) of the Sun with constant L . For small changes ΔT and ΔR with L constant,

$$\frac{\Delta R}{R} = -2 \frac{\Delta T}{T} \quad (1)$$

As $T \sim 6,000$ K, a drop in surface temperature of 1 K would correspond to a relative expansion $\Delta R/R \sim 3 \times 10^{-4}$.

The concept of magnetic buoyancy was introduced by Parker² and Jensen³ and is an important ingredient in the solar dynamo. Consider an isolated magnetic flux tube of field strength B in the solar interior. The sum of the gas pressure p_i and the magnetic pressure $p_m = B^2/8\pi$ inside the tube must balance the external gas pressure p_e , so $p_i + p_m = p_e$, and $p_i < p_e$. If the flux tube is in

thermal equilibrium with its surroundings ($T_i = T_e$), then the density inside the tube is lower than the density of the surroundings ($\rho_i < \rho_e$) and the flux tube is buoyant. Jensen³ pointed out that, because of the density depression inside a magnetic flux tube, the presence of magnetic flux tubes in the solar interior will cause the volume of the Sun to increase; thus, we would expect the radius of the Sun to vary with the solar cycle, with maximum radius occurring near sunspot maximum. Jensen's suggestion, which has attracted little previous attention, is pursued here.

How much expansion and contraction of the Sun might we expect due to variations in magnetic buoyancy over the solar cycle? A precise estimate would require detailed knowledge of the magnetic field distribution in the convection zone over the solar cycle and a complete calculation of the effect of the magnetic field on the structure of the solar envelope. This is beyond present capabilities, so instead a crude calculation is used to get a very rough estimate. Consider a toroidal flux tube that circles the Sun at spherical radius R_1 and constant latitude ϕ . Let the tube have field strength B and cross-sectional radius a . The mass of gas within the tube is less than would occupy the same volume in the absence of the magnetic field. The difference in densities outside and inside the tube is given by

$$\Delta\rho = \rho_e - \rho_i = (p_e - p_i)/RT_1 = p_m/RT_1$$

where $T_1 = T_e = T_i$ is the temperature of the flux tube and its immediate surroundings. The total mass defect in the tube is equal to the volume of the tube times $\Delta\rho$, or

$$\Delta M = 2\pi R_1 \cos \phi \pi a^2 p_m / RT_1$$

Assuming that this mass ΔM is spread over a thin spherical shell of mean radius R_1 and thickness ΔR , then the mass in this shell is given by

$$\Delta M = 4\pi R_1^2 \Delta R \rho_e$$

Equating the two expressions for ΔM gives

$$\frac{\Delta R}{R} = \frac{\pi}{2} \cos \phi \left(\frac{R}{R_1} \right) \left(\frac{a}{R} \right)^2 \frac{p_m}{p_e} \quad (2)$$

This expression gives an estimate of the relative change in radius due to a single toroidal flux tube.

To obtain a numerical estimate we can, for example, use the results of the study by Weiss⁴ of magnetic flux tubes in the solar convection zone, based on equipartition of magnetic energy and kinetic energy of convection. Weiss finds that a typical flux tube has a total flux of 4×10^{21} Mx. The radius and field strength of the flux tube vary slowly from 5,000 km and 5,700 G at the bottom of the convection zone to 14,000 km and 600 G at photosphere. The ratio p_m/p_e varies greatly, from 6×10^{-6} at the bottom of the convection zone to 2.8×10^{-1} at the photosphere. Thus, the dominant contribution to ΔR will come from flux tubes near the top of the convection zone. Using values given by Weiss for a and p_m/p_e in equation (2), for R_1/R in the range 0.98–1.0, we obtain values of $\Delta R/R$ in the range $5 \times 10^{-8} \cos \phi$ to $2 \times 10^{-4} \cos \phi$. For example, a single flux tube at a depth of 1,000 km at middle latitude gives $\Delta R/R \sim 2 \times 10^{-5}$. Now, to estimate the change in solar radius with the solar cycle, we need to estimate the difference in total magnetic flux in the upper convection zone between solar maximum and minimum. If we assume that the total magnetic flux emerging at the solar surface is a good indicator of the total flux in the uppermost layers of the convection zone, then observations suggest a difference in total flux of the order 10^{23} Mx between solar maximum and minimum. This is equivalent to some 25 or more of the individual flux tubes discussed above and could imply a relative change in radius $\Delta R/R \sim 5 \times 10^{-4}$ or more and a corresponding drop in surface temperature of ~ 1.5 K or more.

Magnetic field strengths in the convection zone may be higher than the equipartition limit used in the estimate above. Indeed, observations⁵ have shown that almost all of the magnetic flux in the quiet photosphere is concentrated to strengths of 1,000–2,000 G, well above the equipartition limit, and the theory of

Galloway, Proctor and Weiss⁶ predicts that magnetic fields will be concentrated to 10 times the equipartition limit or more by the convection. Equation (2) can be rewritten as

$$\frac{\Delta R}{R} = \frac{FB \cos \phi}{16\pi R R_1 p_e}$$

where $F = \pi a^2 B$ is the total magnetic flux in the tube. For a tube of fixed total flux, the relative change in radius is proportional to the magnetic field strength. Thus, higher field strengths in the convection zone would increase the estimate of $\Delta R/R$ for a fixed difference in total flux over the solar cycle. With higher field strengths, flux tubes deeper in the convection zone would contribute significantly to $\Delta R/R$. Although there is uncertainty about magnetic fields in the convection zone, the estimates do suggest that the surface cooling of ~ 6 K observed by Livingston may be at least partly caused by expansion due to magnetic buoyancy.

The historical record of measurements of the solar radius gives no clear or consistent picture^{7–9}. The currently accepted value¹⁰ ($R = 959.63''$ at 1 AU) dates back to 1891 and is subject to correction for irradiation¹¹. Variability of R with amplitudes ranging from $\sim 0.05''$ to $\sim 2''$ and periods of 7, 8, 11 and 22 yr has been reported. Giannuzzi¹², for example, found variations in R with amplitude 0.2–0.5'' and period 22 yr. Meyermann¹³ reported variations with total amplitude $\sim 0.15''$ ($\Delta R/R \sim 1.5 \times 10^{-4}$) and period 11 yr with maximum radius occurring very near sunspot maximum; this is in good agreement with the theoretical argument above. Gething⁸ finds variations in R with amplitude ~ 0.1 – $0.2''$, but with no obvious correlation with the solar cycle. Improved measurements of the solar radius over a solar cycle are needed to test the ideas expressed here. Recent improvements^{9,14} in the theoretical definition of the solar radius should give a considerable increase in accuracy. Changes in opacity during expansion may affect the visual change in solar radius. Note that the solar oblateness measured by Dicke and Goldenberg¹⁵ ($\Delta R/R \sim 5 \times 10^{-5}$) and later disputed by Hill and Stebbins¹⁶ is an order of magnitude smaller than the variation in mean radius estimated above. Even if the expansion due to magnetic buoyancy is too small to produce a measurable change in surface temperature, it may still be large enough to show up directly as a change in radius. Accurate monitoring of the solar radius over the solar cycle would give a measure of the variation of total magnetic flux in the convection zone and add to our understanding of the solar dynamo. Measurements of the solar radius from space are particularly desirable.

This research was supported by grants from the US Air Force Geophysics Laboratory and NASA. I thank Alfred Clark, Jr and Eugene Parker for helpful comments.

JOHN H. THOMAS

Department of Mechanical and Aerospace Sciences
and
CEK Mees Observatory,
University of Rochester,
Rochester, New York 14627

Received 6 April; accepted 10 July 1979.

- Livingston, W. C. *Nature* **272**, 340–341 (1978).
- Parker, E. N. *Astrophys. J.* **121**, 491–507 (1955).
- Jensen, E. *Ann. Astrophys.* **18**, 127–140 (1955).
- Weiss, N. O. *Mon. Not. R. astr. Soc.* **128**, 225–235 (1964).
- Harvey, J. W. in *Highlights of Astronomy* (ed. Müller, E.) Vol. 4, 223–239 (Reidel, Dordrecht, 1977).
- Galloway, D. J., Proctor, M. R. E. & Weiss, N. O. *Nature* **266**, 686–689 (1977).
- Goldberg, L. in *The Sun* (ed. Kuiper, G. P.) 1–35 (University of Chicago Press, 1953).
- Gething, P. J. D. *Mon. Not. R. astr. Soc.* **115**, 558–570 (1955).
- Wittmann, A. *Solar Phys.* **29**, 333–340 (1973).
- Auwers, A. *Astr. Nachr.* **128**, 367 (1891).
- Cullen, R. T. *Mon. Not. R. astr. Soc.* **86**, 344–349 (1926).
- Giannuzzi, M. A. *Contr. Sci. Oss. Roma*, No. 211 (1955).
- Meyermann, B. *Astr. Nachr.* **279**, 45–46 (1950).
- Hill, H. A., Stebbins, R. T. & Oleson, J. R. *Astrophys. J.* **200**, 484–498 (1975).
- Dicke, R. H. & Goldenberg, H. M. *Phys. Rev. Lett.* **18**, 313–316 (1967).
- Hill, H. A. & Stebbins, R. T. *Astrophys. J.* **200**, 471–483 (1975).

Silicalite-2, a silica analogue of the aluminosilicate zeolite ZSM-11

THE structure of silicalite, a new polymorph of silica which was first prepared by Flanigen *et al.*¹, is so similar to that of the silica-rich aluminosilicate zeolite ZSM-5 (ref. 2) that we suggest that it should be regarded as the aluminium-free end member of this zeolite series. Therefore it should be possible to prepare silica analogues of other silica-rich aluminosilicate zeolites such as ZSM-11 (ref. 3), ZSM-12 (ref. 4), ZSM-21 (ref. 5) and ZSM-34 (ref. 6). We report here the preparation of what we believe is the analogue of ZSM-11, which we call silicalite-2, adopting the terminology silicalite-1 for the form prepared by Flanigen *et al.*¹. The silicalite-1, ZSM-5, ZSM-11 and, by inference, silicalite-2 frameworks contain 4-, 5- and 6-membered rings linked to form a system of channels with 10-membered ring openings¹⁻³. In silicalite-1 and ZSM-5 the channels are a combination of linear and zig-zag^{1,2}, while in ZSM-11³ and silicalite-2 all the channels are linear. The lack of substitutional aluminium results in silicalites having no significant catalytic or cation exchange properties compared with the zeolites. However, silicalites are unusual in that they are the only known hydrophobic forms of silica and are capable of absorbing organic molecules up to ~6 Å kinetic diameter. Thermogravimetric studies of silicalites-1 and -2 containing absorbed long straight-chain hydrocarbon compounds (C₄–C₁₆) show that degassing occurs in two stages, which we consider to be due to self-blocking of the channels at intersection points.

The precursors of the silicalites are prepared by reacting tetraalkylammonium or tetraalkylphosphonium salts with a reactive form of silica at a high pH, at temperatures between 100 °C and 200 °C, in a closed system^{4,5}. We have found that the silicalite-2 precursor can be prepared using tetra-*n*-butylammonium hydroxide only, although adding ammonium hydroxide or hydrazine hydrate as a source of extra hydroxyl ions increases the reaction rate considerably. A successful preparation was to mix 8.5 mol SiO₂ as silicic acid (74% SiO₂), 1.0 mol tetra-*n*-butylammonium hydroxide, 3.0 mol NH₄OH and 100 mol water in a steel bomb and heat at 170 °C for 3 days. The precursor crystals are ovate in shape, ~2–3 µm long and 1–1.5 µm diameter, similar to those reported for ZSM-11³. The silicalite-2 precursor will not form if Li, Na, K, Rb or Cs ions are present, in which case the silicalite-1 precursor is formed. The size of the tetraalkylammonium ion seems critical because replacement of the tetra-*n*-butylammonium hydroxide by other quaternary ammonium hydroxides (such as tetraethyl, tetrapropyl, triethylpropyl, and triethylbutyl hydroxides) results in amorphous products. Silicalite-2 precursor is stable at extended reaction times in the hydrothermal system. However, in similar conditions the silicalite-1 precursor converts to α-quartz, presumably due to the effect of alkali metal ions. The amount of Al present in silicalite-2 depends on the purity of the starting materials. We typically find <5 p.p.m. Al present. The precursor contains occluded tetraalkylammonium salts which, because of their size, can be removed only by thermal decomposition. We have observed, by means of thermal analysis and mass spectrometry, that the tetraalkylammonium ion decomposes at ~300 °C and is lost as the tertiary amine, alkene and water. This

can be contrasted to the normal thermal decomposition at 200 °C of the same tetraalkylammonium salt in air.

Figure 1 shows the X-ray powder diffraction patterns we have obtained for silicalite-1 and silicalite-2, and for ZSM-5 and ZSM-11. Major differences between the patterns of silicalite-1 and silicalite-2 are that peaks at 9.06, 13.9, 15.5, 16.5, 20.8, 21.7, 22.1, 24.4, 26.6 and 27.0 degrees 2θ (CuKα radiation) are absent from the silicalite-2 pattern. Also, peaks at 8.8, 14.8, 17.6, 23.1, 23.9 and 29.9 degrees 2θ are singlets in the silicalite-2 pattern rather than doublets as in the silicalite-1 pattern. These differences are the same as those found between the diffraction patterns of orthorhombic ZSM-5 and tetragonal ZSM-11. Unit cell dimensions are given in Table 1 calculated on the assumption of tetragonal symmetry for silicalite-2, and show that the effect of the substituted aluminium in ZSM-5 and ZSM-11 is to increase the unit cell size over that of silicalite-1 and silicalite-2. The measured densities and refractive indices of silicalite-2 and its precursor are 1.82 and 1.98 g cm⁻³ and 1.41 and 1.48 respectively. Calculation of the unit cell volumes shows that silicalite-2 should be ~1% less dense than silicalite-1, whereas we find that silicalite-2 is ~3% more dense. The samples of silicalite-2 prepared here possibly contain a small percentage of another 'more dense' crystallographic form of silica.

Desorption or degassing of various organic molecules from silicalite-1 and silicalite-2 has been studied by thermogravimetric analysis. Rigid molecules, such as *p*-xylene, are lost in a single low-temperature step due to simple desorption. However, flexible molecules, such as *n*-alkanes or *n*-alcohols of chain length >C₅ are lost in two steps (Fig. 2). We believe this shows that self-blocking of the channel structure is occurring, probably by molecules bending through 90° into an intersecting channel,

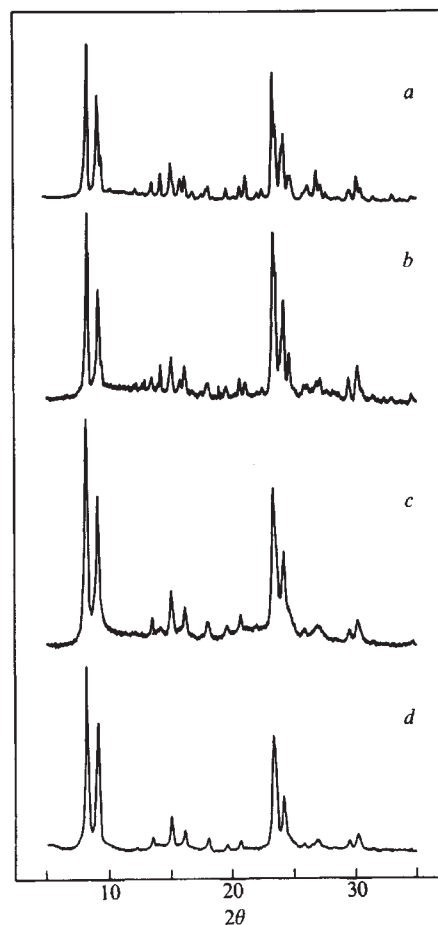


Fig. 1 X-ray diffraction patterns of silicalite-1 (a) and silicalite-2 (c) and the corresponding aluminosilicate zeolites ZSM-5 (b) and ZSM-11 (d).

Table 1 Lattice constants

	a	b	c
Silicalite-1	20.07 (1)	19.86 (1)	13.36 (1)
ZSM-5 ²	20.1	19.9	13.4
Silicalite-2	20.04 (1)	20.04 (1)	13.38 (1)
ZSM-11 ³	20.12	20.12	13.44

The figures in parentheses are 1 s.d.

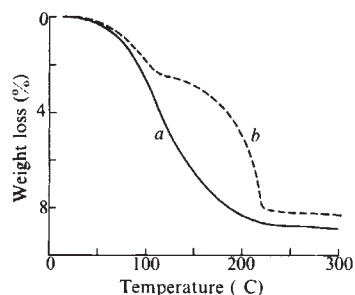


Fig. 2 Thermogravimetric curves showing the weight loss of: *a*, *p*-xylene; and *b* decane from silicalite-2 (heating rate = $2.5^{\circ}\text{C min}^{-1}$, atmosphere is air).

and thus preventing egress of other molecules from both of the channels. A higher activation energy is required to overcome the awkward passage of these molecules, resulting in a higher temperature for weight loss when compared with desorption straight through the channel system. Initial results show that there is less self-blocking in silicalite-2 than in silicalite-1. Presumably the zig-zag channels in silicalite-1 are more prone to self-blocking since molecules are required to change direction even when continuing along the same channel, thus increasing the probability of bending into an intersecting channel.

D. M. BIBBY
N. B. MILESTONE
L. P. ALDRIDGE

Chemistry Division, DSIR,
Private Bag, Petone, New Zealand

Received 24 May; accepted 21 June 1979.

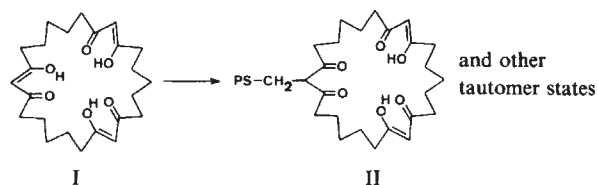
1. Flanigen, E. M. *et al.* *Nature* **271**, 512 (1978).
2. Kokotailo, G. T., Lawton, S. L., Olson, D. H. & Meier, W. M. *Nature* **272**, 437 (1978).
3. Kokotailo, G. T., Chu, P., Lawton, S. L. & Meier, W. M. *Nature* **275**, 119 (1978).
4. US Pat. No. 3832449 (1974).
5. US Pat. No. 4046859 (1977).
6. US Pat. No. 4079096 (1978).

Extraction of uranium from seawater by polymer-bound macrocyclic hexaketone

THE energy crisis means that a petroleum substitute is urgently needed and a promising candidate is nuclear energy. Uranium, however, like petroleum, is distributed in small limited areas. These circumstances could be dramatically improved if uranium in seawater were to be utilised as a resource. Although its concentration is extremely low (3.3 parts per 10^9), the total amount is huge, calculated at $4\text{--}4.5 \times 10^9$ tons. We have recently synthesised a macrocyclic hexaketone (I) which is a very strong host of uranyl (UO_2^{2+}) ion¹. Its noteworthy characteristics are: (1) six oxygen atoms can be directed towards the inside of the ring, if necessary, to form hexadentate coordination in near coplanarity which corresponds to the crystalline structure of various uranyl salts^{2,3}. (2) β -Diketone can easily be dissociated in seawater ($\text{p}K_a$ is 8.95 for acetylacetone) to form a strong ligand, ketoenolate anion⁴. (3) The bound uranyl ion is readily liberated by the treatment with dilute aqueous acid. Here we report the successful extraction of uranyl ion directly from seawater using polymer-bound macrocyclic hexaketone.

The hexaketone (170 mg, 0.4 mmol) was treated with partially chloromethylated polystyrene (chloromethyl content: 18.8%, 82 mg, 0.14 mmol equiv. for CH_2Cl unit) in dimethylformamide (1 ml) in the presence of anhydrous potassium carbonate (138 mg, 1.0 mmol). An IR spectrum of the resulting polymer, after the usual purification, showed that the CH_2Cl

absorption had almost disappeared and new absorptions at $1,720$ and $1,600\text{ cm}^{-1}$ had appeared which correspond to keto and enol forms of hexaketone unit.



Polymer-bound hexaketone (II) (98 mg, 0.1 mmol equiv. of (I)) dissolved in chloroform (30 ml) was stirred magnetically with 3.5 l of seawater sampled at 20 m off the coast of Ohara, Sotoboh, Chiba Prefecture, Japan (Fig. 2). After 18 h, a chloroform solution was separated and seawater was extracted with chloroform ($5\text{ ml} \times 2$) to recover any dissolved polymer and/or polymer-uranyl complexes. From the combined chloroform solution, uranyl ion was liberated by the treatment with 1 M HCl solution (5 ml)¹. The resulting acid solution was analysed by the arsenazo III method⁵, showing that 0.84 p.p.m. of uranyl ion is present which corresponds to 36% of the UO_2^{2+} present in the original seawater. An independent analysis of UO_2^{2+} by atomic absorption spectrophotometry was also made to demonstrate the presence of approximately the same amount of uranyl ion. Repeated (three times) extraction and liberation of uranyl ion using the recovered hexaketone (II), accumulated 1.5 p.p.m. of UO_2^{2+} in 9.2 ml of acid solution ($13.8\text{ }\mu\text{g}$) (total seawater treated is $3 \times 5\text{ l}$). After these cycles, the same procedure using the hexaketone (II) recovered, was repeated to extract $11\text{ }\mu\text{g}$ of UO_2^{2+} from further $3 \times 5\text{ l}$ of sea water. This demonstrated that the recovery and regeneration of hexaketone (II) was satisfactory. The extraction time was then reduced to 1 h and gave practically the same efficiency. These results clearly indicate that the polymer-bound hexaketone (II) can successfully extract uranyl ion directly from seawater and be liberated in concentrated form on treating with a dilute aqueous acid solution.

Attempts to extract uranium from seawater have so far concentrated on the method using the adsorbent hydrous titania sorber⁶. However, economical and environmental analyses are discouraging at present; the major problem is the large loss of titania sorber. It has been estimated that 68 pounds are needed to produce 1 pound of uranium⁷. The method described here avoids an appreciable loss of adsorbent by cross linking the supporting material via chemical bonding, and on further refinements it will give adsorbents specific to uranyl ion. The present procedure, however, is still not practical because the polystyrene backbone is too hydrophobic to achieve an efficient rapid contact with seawater if used in powdered form, although slow extraction does take place. In our preliminary experiments using powder of a partially sulphonated polystyrene, however,

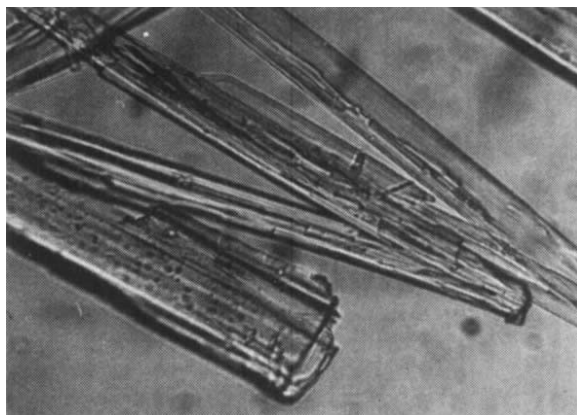


Fig. 1 Light microscopy of crystals of the hexaketone (I).

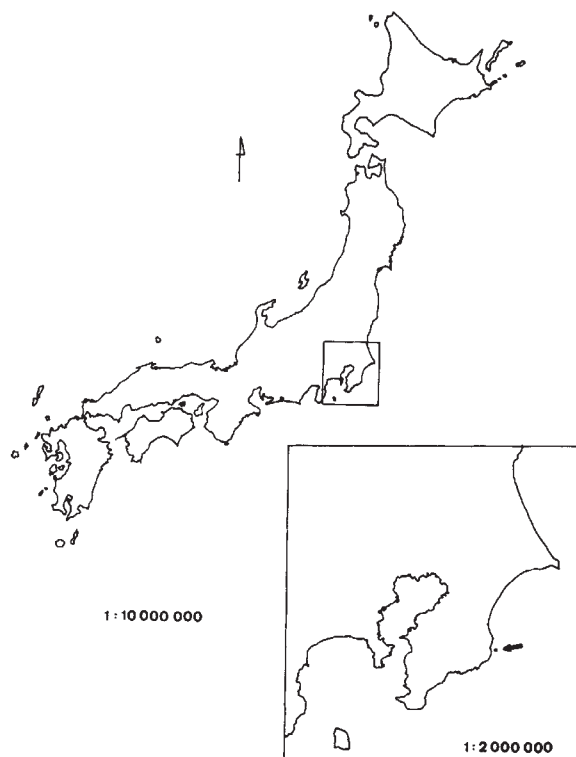


Fig. 2 Seawater was sampled at the spot asterisked along the Black Current.

uranyl ion was very effectively and rapidly extracted from seawater. The polystyrene was partially sulphonated by chlorosulphonic acid followed by the partial chloromethylation. Hexaketone was then attached to this polymer as described above. The details will be published elsewhere.

We thank Idemitsu Kosan Co. Ltd. for the supply of seawater.

IWAO TABUSHI
YOSHIKI KOBUE
TAKAKO NISHIYA

Department of Synthetic Chemistry,
Faculty of Engineering,
Kyoto University,
Yoshida, Kyoto 606, Japan

Received 23 April; accepted 10 July 1979.

1. Tabushi, I., Kobue, Y. & Nishiyama, T. *Tetrahedron Lett.* (in the press).
2. Fankuchen, I. *Z. Kristallogr.* **91**, 473–479 (1935).
3. Zachariasen, W. H. *Acta crystallogr.* **1**, 277–281, 281–285 (1948).
4. Mehrotra, R. C., Bohra, R. & Gaur, D. P., in *Metal β-Diketones and Allied Derivatives* (Academic, New York, 1978).
5. Onishi, H. & Toita, Y. *Bunseki Kagaku* **14**, 1141–1146 (1965).
6. Davies, R. V., Kennedy, J., McIlroy, R. W., Spence, R. & Hill, K. M. *Nature* **203**, 1110–1115 (1964).
7. Harrington, F. E. *et al.* Oak Ridge Nat. Lab., ORNL TM-4757, 1–77 (1974).

Anodic coal reaction lowers energy consumption of metal electrowinning

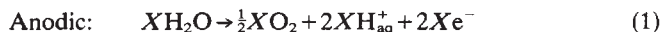
THE metals which can be extracted in large amounts from aqueous solutions of their salts by electrowinning are Cr, Mn, Co, Ni, Cu, Zn, Ga, Cd, In and Tl. Gilroy¹ specifies an annual electrowinning production for 1971 of 2,400,000 tons of Zn and 540,000 tons of Ca, 53% and 10% respectively of total production at that time. We describe here a modification to the established electrowinning technique which lowers the energy consumption of the process.

Table 1 Current efficiency and electrical energy consumption during copper deposition in the conventional process

Reaction rate I (mA)	Total coulomb consumed It (C)	Theoretical Cu equivalent to the charge passed W_1 (g)	Measured amount of Cu deposited W_2 (g)	Current efficiency $(W_2/W_1) \times 100$ (%)	Electrical power consumed $E_1 = (I \int V dt)/W_2$ (kWh kg ⁻¹)
5.90	100.79	0.03319	0.03027	91.2	1.52
7.95	147.39	0.04853	0.04620	95.2	1.51
10.00	144.00	0.04741	0.04425	93.3	1.52
12.00	133.90	0.0442	1.04288	97.2	1.622

Electrolyte, 0.125 M CuSO₄ and 0.5 M H₂SO₄; Temperature, 60 °C.

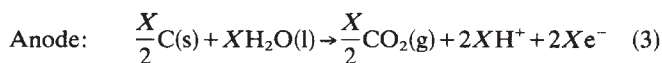
The reactions for conventional electrowinning of a metal M are:



The reversible thermodynamic potential, E° , of reaction (1) is 1.229 V. It and associated polarisation overpotentials, E_{op} , constitute a large portion of the total potential E_{tot} of the practical electrowinning cell:

$$E_{\text{tot}} = E^\circ + E_{\text{op}} + E_m$$

where E_m is the half-cell potential for reducing the metal ion, reaction (2). Browall *et al.*² introduced the reducing agent, CO, at the anode to lower the cell potential for electrolytic production of hydrogen. Using coal as the reducing agent, we devised^{3–5} a new electrochemical process which converts coal and water into two separate gaseous products—one at the anode, comprising essentially gaseous oxides of carbon, and the other at the cathode which is essentially pure hydrogen. We called this process electrochemical gasification of coal or, alternatively, coal-assisted water electrolysis. Coal contains a wide variety of organic molecules rich in carbon. Focusing only on the carbon of the coal the reactions of the electrochemical coal-gasification process, in simplified stoichiometry, would be:



The reversible thermodynamic potential, E° , of half-cell reaction (3) is only about 0.21 V. We have now extended electrochemical coal gasification to the electrowinning of metals by substituting the half-cell reaction (3) (which consumes coal) for the half-cell reaction (1) which usually takes place during conventional electrowinning of metals from aqueous solutions of

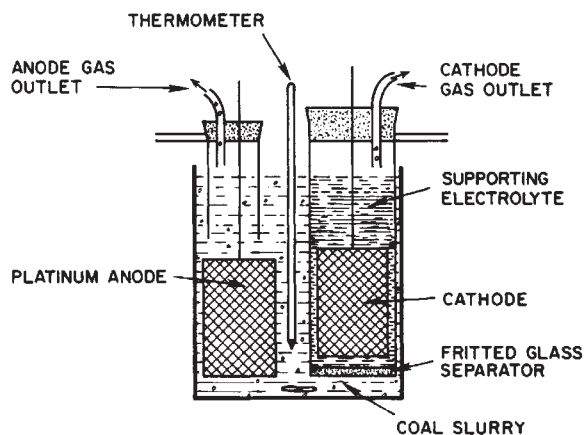


Fig. 1 Coal gasification cell: electrolyte volume, 80 cm³; anode, 5.3 cm² (geometrical) platinum mesh, gauge 52; cathode, 8.4 cm² (geometrical) platinum mesh, gauge 52; stirring, magnetic, 7/8-inch magnet).

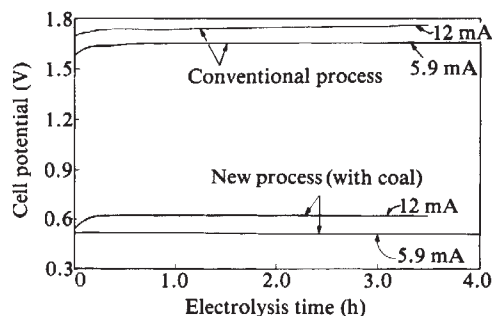


Fig. 2 Comparison of cell potentials of the copper electrowinning processes in galvanostatic conditions: Electrolyte, 0.125 M CuSO_4 and 0.5 M H_2SO_4 ; coal sample, N. Dakota lignite; slurry concentration, 0.15 g cm^{-3} ; particle size, 74–88 μm ; temperature, 60°C .

their salts. By this coal-based innovation the total overall cell potential of the resulting metal electrowinning process is lowered by about 1.0 V with a corresponding significant reduction in the consumption of electrical energy. This new method of electrowinning from aqueous electrolytes is illustrated using preliminary experimental results for the coal-assisted copper electrowinning process.

Experiments were conducted in the simple cell shown in Fig. 1. The external e.m.f. was controlled by a potentiostat (PAR 371) and the electrodes (both anode and cathode) were Pt mesh, gauge 52 (0.004-inch diameter wire, Mathey Bishop Inc). Aqueous electrolytes were prepared with Baker analysed reagent grade sulphuric acid and Baker $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ crystals. Coal slurries were prepared from pulverised North Dakota lignite coal; the ultimate analyses are described elsewhere⁴. Metal analysis was carried out on a Perkin-Elmer 460 atomic absorption spectrophotometer.

For comparison Cu was first deposited by conventional electrowinning without using coal. Accordingly, using the cell of Fig. 1, Cu was deposited on a platinum-mesh cathode from aqueous electrolyte (0.125 M CuSO_4 in 0.5 M H_2SO_4) at 60°C , and at four different galvanostatic currents: 5.9, 7.95, 10 and 12 mA. The corresponding change in cell potential was recorded and typical cell potentials corresponding to the galvanostatic currents of 5.9 and 12 mA are plotted in Fig. 2. The conventional copper electrowinning reaction took place at corresponding cell potentials of about 1.65 and 1.73 V respectively. Figure 2 also shows the results obtained in parallel experiments, identical except that the coal was simultaneously oxidised at the anode according to equation (3) while Cu cations were reduced at the cathode according to equation (2). In these latter experiments, North Dakota lignite (NDL) coal slurry (0.15 g cm^{-3}) was introduced into the anode compartment of the cell and was anodically oxidised while copper deposited on the cathode in identical galvanostatic experimental conditions. As a result of the oxidation or gasification of the coal at the anode, the overall cell potential was lowered by $\sim 1.1 \text{ V}$, compared to the conventional process. This is evident from Fig. 2. It is seen that the simultaneous anodic oxidation of coal lowers the potential by $\sim 1.10 \text{ V}$, a value which is very close to the difference in half-cell potential between reaction (1) and (3).

Data on the current, electricity consumption and weight of Cu deposited are given in Table 1 for the conventional process, and in Table 2 for the new, coal-assisted process. Current efficiencies are also given computed from the ratio of the amount of Cu actually deposited to the equivalents of the charge passed during the experiment. The introduction of coal into the anode compartment during electrolysis evidently reduces current efficiency only slightly, at the worst about 2%. Comparing the right-hand columns of Tables 1 and 2 reveals that the simultaneous anodic oxidation of coal during Cu electrowinning reduces the electrical energy consumption by a factor of about two-thirds or more. Whereas the conventional process uses only electrical energy, our new process uses both electrical energy

(E_2) and the chemical energy (E'_3) of coal directly. For a reasonable comparison of the energy consumption of the two processes, the chemical energy of the coal consumed should be converted into equivalent electrical energy. By the stoichiometry of reaction (3), the mass of carbon consumed to produce 1 kg of copper is approximately:

$$\frac{12 \text{ kg C}}{\text{kg atom C}} \times \frac{1 \text{ kg atom C}}{2 \text{ kg atom Cu}} \times \frac{1}{63.54} \frac{\text{kg atom Cu}}{\text{kg Cu}} = 0.0944 \frac{\text{kg C}}{\text{kg Cu}}$$

Based on the enthalpy of combustion of carbon, this corresponds to a chemicothermal energy consumption of

$$E'_3 = 0.0944 \frac{\text{kg C}}{\text{kg Cu}} \times \frac{94,052 \text{ kcal}}{12 \text{ kg C}} \times \frac{4180 \text{ J}}{\text{kcal}} \times \frac{1 \text{ kW s}}{10^3 \text{ J}} \times \frac{1 \text{ h}}{3,600 \text{ s}} \\ = 0.859 \frac{\text{kWh}}{\text{kg Cu}}$$

Assuming a practical efficiency of 35% for conversion of the combustion enthalpy of carbon (coal) into electrical energy, the electrical energy equivalent to the coal consumed is:

$$E_3 = (0.859)(0.35) = 0.300 \frac{\text{kWh}}{\text{kg Cu}}$$

Using these values of E'_3 and E_3 with the values of E_1 and E_2 from Tables 1 and 2, the total energy requirements and also the net energy savings have been calculated for the new process and are reported in Table 3. Compared to the conventional process, the new coal-consuming process requires about 45% less net equivalent electrical energy for producing the same amount of copper.

A more telling comparison of these processes might be on the basis of relative energy costs based on assumed costs of US \$0.02 per kWh for electricity and \$35.00 per ton (\$0.0514 per kg C) for coal assumed to be 75% carbon. Using mean values for E_1 and E_2 these costs are:

For the conventional process: $\text{Cost} = (\$0.02) \cdot \bar{E}_1 = \0.031 per kg Cu.

For the new coal-assisted process: $\text{Cost} = (\$0.02) \cdot \bar{E}_2 + (0.0944)(\$0.0514) = \$0.0147$ per kg Cu.

The savings in energy cost appears, therefore, to be $\sim 50\%$.

In the experiments discussed so far, NDL coal which contains about 11.7% ash has been used as a reactant. The ash composition suggests the possibility of the presence of impurities such as Al, Fe, Ti, Ca, Mg, Na and K ions in the electrolytes. Consideration of the thermodynamic potential of the $\text{M}-\text{M}^{x+}$ couple of these metals indicates that only iron might interfere with the copper deposition for which half-cell potentials have been reported⁶ as follows:

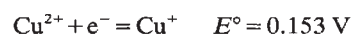
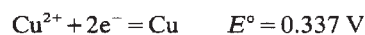


Table 2 Current efficiency and electrical energy consumption during copper deposition: new coal-assisted process

Reaction rate I (mA)	Total coulomb consumed It (C)	Theoretical Cu equivalent to the charge passed W_1 (g)	Measured amount of Cu deposited W_2 (g)	Current efficiency (W_2/W_1) $\times 100$ (%)	Electrical power consumed $E_2 = (I \int V dt)/W_2$ (kWh kg ⁻¹)
5.90	105.84	0.03485	0.031115	89.4	0.468
7.95	132.61	0.04366	0.04062	93.0	0.428
10.00	142.80	0.04702	0.04407	93.3	0.524
12.00	129.60	0.04267	0.04063	95.2	0.552

Electrolyte, 0.125 M CuSO_4 and 0.5 M H_2SO_4 ; Temperature, 60°C ; North Dakota lignite coal concentration 0.15 g cm^{-3} ; particle size, 74–88 μm .

Table 3 Savings in energy consumption

Galvano- static reaction rate I (mA)	Total energy consumption in the new process $E_2 + E_3$ (kWh per kg Cu)	Net equivalent electrical energy consumed in this new process $E_2 + E_3$ (kWh per kg Cu)	Net savings in equivalent electrical energy $E_1 - E_2 - E_3$ (kWh per kg Cu)	Net equivalent energy savings ($E_1 - E_2 - E_3$) E_1 $\times 100$ (%)
5.9	1.328	0.77	0.75	49.3
7.95	1.29	0.73	0.78	51.6
10.0	1.384	0.826	0.694	45.7
12.0	1.412	0.852	0.768	47.3

Reported⁶ potentials of the different iron couples are:



The ash of NDL coal contains 12.5% Fe_2O_3 (other coals contain even less iron). Therefore iron is the only large-scale impurity which should be removed in a continuous process because reduction from the ferric to the ferrous state would reduce current efficiency for copper deposition and would increase the power consumption. Note that iron is generally also the only large-scale impurity present in copper ore itself and the technology for iron removal is well established, for example, by air oxidation of Fe^{2+} to Fe^{3+} which can easily be made to precipitate as the hydroxide. In some industrial processes ferric iron concentration is kept high to promote faster leaching of sulphide ores.

The purity of the electrodeposited copper was determined by dissolution in HNO_3 followed by atomic absorption spectrophotometry. In all experiments the purity of the deposited copper was >99%. The use of NDL or the presence of its ash in the electrolyte does not seem to have any significant effect on the copper purity. While our results, showing significant reduction in the energy requirement, pertain to one specific reaction system, similar observations are expected for other metal electrowinning reactions in aqueous electrolytes. Anode materials other than platinum have also been investigated. Graphite rods and graphite felt anode (Union Carbide Corp.) performed as well as platinum; copper plate cathodes have also been successfully tested.

Although, in an energy conscious environment this method looks promising, the practical applicability of the process depends on: (1) Achieving current densities sufficiently large to keep the capital cost economical; in our simple laboratory experiments no attempt has been made to increase current densities which were far lower than those used in industrial processes. (2) Almost total consumption of the carbon-value of the coal particles. (3) The influence of the particular coal and its ash content on electrowinning kinetics and on the properties and morphology of the deposited metal. (4) Preventing build-up in a continuous process of impurity cations which might compete with copper ions in electrolysis; such impurities would have to be removed continuously, for example by air oxidation of Fe^{2+} to precipitate ferric hydroxide.

This research was supported by the University of Connecticut Research Foundation and the US Department of Energy. We thank Larry Veneziano for technical assistance.

MOHAMMAD FAROOQUE
ROBERT W. COUGHLIN

Department of Chemical Engineering,
University of Connecticut,
Storrs, Connecticut 06268

Received 14 May; accepted 4 July 1979.

1. Gilroy, D. *The Electrowinning of Metals, Industrial Electrochemical Processes* (ed. Kuhn, A.) Ch. 6 (Elsevier, New York, 1971).
2. Browall, K. W. & Hanneman, R. E. *General Electric Technical Information Ser., Rep. No. 75CRD012* (1975).
3. Coughlin, R. W. & Farooque, M. *Nature* **279**, 301–303 (1979).
4. Coughlin, R. W. & Farooque, M. (IC & E).
5. Farooque, M. & Coughlin, R. W. *Fuel* (in the press).
6. Latimer, W. M. *Oxidation Potentials* (Prentice-Hall, Englewood Cliffs, 1952).

Impact of CO_2 on cooling of snow and water surfaces

THE levels of CO_2 in the atmosphere are being increased by the burning of fossil fuels and reduction of biomass^{1–3}. It has been calculated that the increase in CO_2 levels should lead to global warming because of increased absorption by the atmosphere of terrestrial longwave radiation in the far IR ($>5 \mu\text{m}$)^{4–7}. From model computations, CO_2 is expected to produce the largest climatic effect in high latitudes by reducing the size of ice and snow fields⁴. We present here computations of spectral radiative transfer and scattering within a snow pack and water. The results suggest that CO_2 significantly reduces the shortwave energy absorbed by the surface of snow and water. The energy deficit, when not compensated by downward atmospheric radiation, may delay the recrystallisation of snow and dissipation of pack-ice and result in a cooling rather than a warming effect.

We consider here the spectral absorption of CO_2 in the near IR portion (wavelength $0.75 \approx 5.0 \mu\text{m}$) of incoming shortwave radiation with respect to the role of near IR in the heating of the uppermost (active) layer of snow and of water. Calculations were made to determine the impact of doubled CO_2 on the amount of short wave radiation absorbed in these layers in conditions typical of the start and the end of the local snow and ice season^{8–10}. At this stage, other components of the heat balance such as transfer of sensible and latent heat, longwave atmospheric radiation, heat convection, and so on, were not considered.

To study the solar heating of snow, analytic expressions for the flux and divergence of flux (heat flux) were obtained by solving the radiative transfer equation in the modified two-stream approximation¹¹. The scattering and absorption within the snow-cover by the randomly distributed ice grains were represented by a single scattering albedo and anisotropic phase function. Geometric optics calculations^{11,12} were used to relate the scattering and absorption parameters to the grain size and density. The spectral flux, $F_\lambda(z)$, and the negative of the derivative of spectral flux, $F^1_\lambda(z)$, at a depth z within the snowpack were obtained from the equations:

$$F_\lambda(z) = [I_+(z) - I_-(z)] f_\lambda W_0$$

$$F^1_\lambda(z) = \sqrt{2} \gamma_e (1 - \omega) [I_+(z) + I_-(z)] f_\lambda W_0$$

where

$$I_+(z) = c_1 \left[1 + \frac{(1 - \omega)\sqrt{3}}{a} \right] \exp(-a\tau) + c_2 \left[1 - \frac{(1 - 2)\sqrt{3}}{a} \right] \exp(a\tau)$$

$$I_-(z) = c_1 \left[1 - \frac{(1 - \omega)\sqrt{3}}{a} \right] \exp(-a\tau) + c_2 \left[1 + \frac{(1 - \omega)\sqrt{3}}{a} \right] \exp(a\tau)$$

$$c_1 = \frac{1}{\left[1 + \frac{\sqrt{3}(1 - \omega)}{a} \right] \left[1 + \frac{\alpha(\alpha - R)}{R\alpha - 1} \exp(-2a\tau_0) \right]}$$

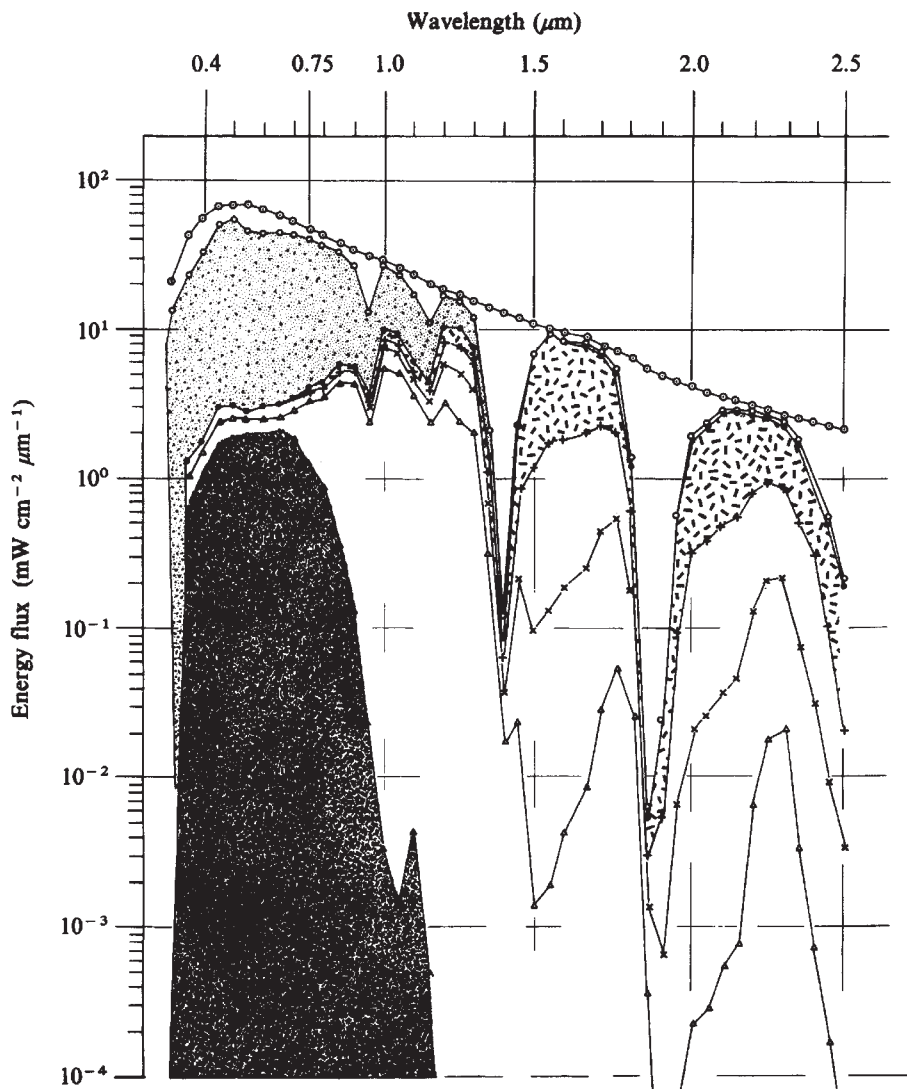


Fig. 1 Shortwave radiation flux in a snow pack in clear dry atmosphere at site A (see text), with 23° solar elevation and normal CO₂ level. Reflected radiation light stippled; energy absorbed in the uppermost 0.2 cm of a snowpack is marked with an intermediate tone; and energy penetrating into the underlying surface is heavy dark stippled. Almost all energy absorbed by snow is in the wavelength above 0.75 μm. ○ Top atmosphere; ○ surface; ● surface penetration; + 0.2 cm; × 0.5 cm; △ 1 cm; ▽ 15 cm.

$$c_2 = \frac{\left(\frac{\alpha - R}{R\alpha - 1} \right) \exp(-2a\tau_0)}{\left[1 + \frac{\sqrt{3}(1-\omega)}{a} \right] \left[1 + \frac{\alpha(\alpha - R)}{R\alpha - 1} \exp(-2a\tau_0) \right]}$$

where $I_+(z)$ is the intensity of downward oriented radiation; $I_-(z)$ the intensity of upward oriented radiation; f_λ the fraction of the solar flux incident on the snowpack within a narrow band centred around wavelength λ ; W_0 the insolation to the top of the atmosphere; ν the radius of ice grains (cm); ρ the density of the snowpack (g cm⁻³); h the thickness of the snowpack (cm); β the fraction of energy scattered in the backward hemisphere in single scattering (≈ 0.065); κ_λ the spectral absorption coefficient of ice (cm⁻¹); $C_{1,2}$ the integration constants; R the reflectance of the underlying snow surface (ice); $\tau = \gamma_e z$ the optical thickness; and $\tau_0 = \gamma_e h$ the total optical thickness of the pack.

Also

$$a = (3(1-\omega)(1-\omega+2\omega\beta))^{1/2}$$

$$\alpha = 1 - \frac{2(1-\omega)^{1/2}}{(1-\omega+2\omega\beta)^{1/2} + (1-\omega)^{1/2}} = \text{spectral reflectance}$$

$$\gamma_e = \left(\frac{3}{2\nu} \right) \frac{\rho}{0.917} = \text{extinction coefficient}$$

$$\omega = \frac{1}{2} + \frac{1}{2} \exp(-1.67\kappa_\lambda\nu) = \text{single scattering albedo}$$

In the absence of thermal conduction the rate of radiative heating was given by the equation:

$$H = \frac{1}{c\rho} \sum_{\lambda} F_{\lambda}^1(z)$$

where $c = 0.5$ is the specific heat of snow.

The values of the spectral reflectance, and the asymptotic flux extinction coefficient calculated from the model¹³ are in good agreement with the observations¹⁴. We used the Delta-Edington approximation to calculate the radiative flux transfer and absorption in water¹⁵; the albedo was computed after Ivanoff¹⁶. Results agree with observations¹⁷. Atmospheric and radiative conditions approximately represent two sites. Site A is on land at latitude 43°N; water content in the atmospheric column is 0.9 g cm⁻² and the Ångström turbidity factor 0.08. Noon solar elevation is 23°, as in the second half of December. Snow is underlain by soil. Site B is at latitude 78°N; atmospheric water content is 0.5 g cm⁻² and the Ångström turbidity factor 0.02. Noon solar angle is 35°, as in early June when snow melt starts on top of the sea-ice.

In both cases the ozone path length is 0.35 cm, barometric pressure is 1,000 mbar, snow density is 0.22 g cm⁻³, grain radius is 0.35 mm (that of moderately aged snow) and snow thickness is 15 and 30 cm. Normal concentration of CO₂ was conservatively taken as 315 p.p.m. although considerably higher winter values have been reported from the Northern Hemisphere in recent years¹⁸. Computations were performed for solar elevation changed in a daily cycle at hourly intervals. Flux was computed at a depth of 0, 0.2, 0.5, 1.0, 2.0, 5.0, 15.0 and 30.0 cm for snow

Table 1 Shortwave generated heating in snow and water

Medium	Case	Solar Time	Solar elevation (deg)	CO ₂	Heating rate (°C min ⁻¹) depth (cm)						
					0.0	0.2	0.5	1.0	5.0	15.0	100.0
Snow	A	9	11.46	Normal	1.7989	0.5441	0.1738	0.0806	0.0092	0.0002	
				Double	1.7570	0.5363	0.1731	0.0806	0.0092	0.0002	
				Diff. %	2.3292	1.4336	0.4028	0.0000	0.0000	0.0000	
		12	23.00	Normal	4.1561	1.2669	0.4111	0.1922	0.0226	0.0005	
				Double	4.0772	1.2523	0.4097	0.1922	0.0226	0.0005	
				Diff. %	1.8984	1.1524	0.3406	0.0000	0.0000	0.0000	
	B	1	5.48	Normal	0.9277	0.2841	0.0925	0.0433	0.0050	0.0005	
				Double	0.9033	0.2795	0.0920	0.0433	0.0050	0.0005	
				Diff. %	2.6302	1.6192	0.5405	0.0000	0.0000	0.0000	
		7	23.16	Normal	4.6027	1.4188	0.4664	0.2180	0.0261	0.0026	
				Double	4.5121	1.4018	0.4647	0.2180	0.0261	0.0026	
				Diff. %	1.9684	1.1982	0.3645	0.0000	0.0000	0.0000	
Snow	B	12	35.00	Normal	7.0564	2.1792	0.7167	0.3333	0.0399	0.0040	
				Double	6.9340	2.1561	0.7144	0.3333	0.0399	0.0040	
				Diff. %	1.7346	1.0600	0.3209	0.0000	0.0000	0.0000	
	A	12	23.00	Normal	1.1055	0.0984	0.0465	0.0288	0.0068	—	0.0004
				Double	1.0564	0.0981	0.0464	0.0288	0.0068	—	0.0004
				Diff. %	4.4414	0.2946	0.0646	0.0000	0.0000	—	0.0000
Water	A	12	23.00	Normal	1.1055	0.0984	0.0465	0.0288	0.0068	—	0.0004
				Double	1.0564	0.0981	0.0464	0.0288	0.0068	—	0.0004
				Diff. %	4.4414	0.2946	0.0646	0.0000	0.0000	—	0.0000

See text for atmospheric composition and snow properties.

and an additional 100 cm for water. In each case, parallel computation was done for a normal atmospheric content of CO₂, and then for the doubled level. The attenuation spectra of CO₂, water vapour, ozone and aerosols were taken from refs 19 and 20. CO₂ absorbs in the near IR bands 1.40–1.65 μm , 1.90–2.10 μm , and 2.80–2.90 μm .

Figure 1 shows the calculated spectral distribution of the shortwave radiation flux at different depth levels of the snow-pack. The near IR is mostly absorbed in the upper 0.5 cm whereas the visible light is reflected or penetrates to the bottom of a pack. Similarly, in clear freshwater, the near IR is absorbed at the surface, whereas the visible radiation penetrates to deeper layers. Over 90% of the shortwave radiation absorbed in the uppermost 2 mm of snow and water is in the spectral range greater than 1 μm (Fig. 2).

Doubled CO₂ cuts the shortwave-generated heating rate at the snow surface on sites A and B by about 2% (1.7–2.6%, see Table 1) and that of the water surface on site A at noon by 4.4%. At a depth of 0.2 cm the difference is 1.0–1.6% in snow, but only 0.3% in water. The heat flux in snow and water layers below 1 cm is not affected.

Establishment of snow cover profoundly alters the radiative exchange taking place at the land surface because of the high visible albedo and high thermal emissivity of snow. Snow fields become areas of intensive radiative cooling where cold and dry polar atmospheric masses develop. A few days after a snowfall both the albedo and the surface emissivity of snow decreases. This is mainly due to the recrystallisation of snow in the active layer. The changes take place even in uninterrupted subfreezing weather^{21–23}. The energy for snow metamorphism is provided largely from the absorbed near IR. As the shortwave transmissivity of ageing snow increases, the temperature at the base of the pack can rise to a sufficiently high level and initiate the bottom melt. Weaver *et al.*²⁴ described such internal melt of snow in subfreezing weather on top of the fast ice in the Western Davis Strait. Incipient meltwater puddles formed on the ice-snow interface. During 1971–73 the process began in early June and lasted 2–3 weeks, before the snow dissipated and open water pools formed.

It is logical to expect that the CO₂ related reduction of the near IR to the snow surface would delay the recrystallisation, maintain high albedo, and consequently delay the onset of subsurface melt.

Arctic sea-ice breaks under wind stress throughout the year,

exposing leads of open water. In winter when the radiative input is low, and evaporative cooling strong, the water refreezes and the cracks are sealed. In spring and summer the exposed water does not refreeze, the ice floes disintegrate by wave action and are blown away into the open water where they melt. Thus the break-up starts when the absorbed radiation in the leads is sufficiently high to prevent refreezing. In clear cold weather, the

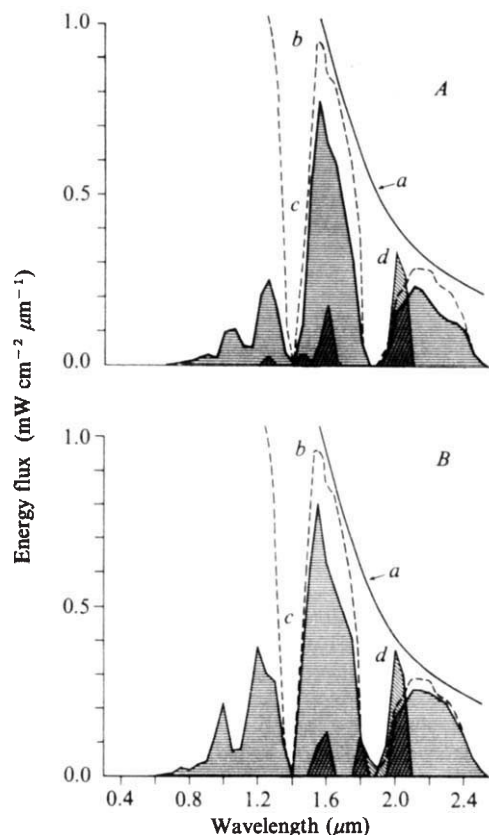


Fig. 2 Shortwave radiation absorbed in the uppermost 0.2 cm of snow (A) and clear fresh water (B) on site A at noon (see text). a, radiation at the top of the atmosphere; b, radiation reaching the surface; c, absorbed radiation; d, deficit due to CO₂ doubling, $\times 10$.

increased CO₂ is likely to delay the onset of the break-up. It is also likely to facilitate an earlier freeze-up of shallow land based fresh water bodies. Apart from CO₂, some other trace gases such as NO₂ or NH₄, and the aerosols of volcanic or industrial origin also absorb in the critical bands and can have a similar effect.

We have to stress that several important radiative and micrometeorological processes were omitted in our model. For instance, the added heating of the atmosphere due to the presence of CO₂ was not calculated. The compensation cannot be complete in a dry subfreezing air of polar origin²⁷. But in summer, when the air over the pack-ice in the central Arctic is relatively moist and thick clouds develop, less near IR reaches the ground, and CO₂ probably warms the system.

Obviously we are far from a comprehensive determination of the effect of CO₂ doubling on snow and ice fields. Our study demonstrates, nevertheless, that the final assessment of the CO₂ impact on global climate is too complicated to be reliably solved by simple annually-averaged general circulation models. Predictions of global warming derived from them should be considered highly tentative at best²⁷. For reliable answers, improved models will have to be developed to handle in detail the different heating response of various surface types to changing spectral irradiance. Field observations will be needed to test the models.

The global impact of CO₂ increase remains uncertain, but it is possible that the related reduction of shortwave radiation absorbed in snow active layers and at the water surface may contribute to an extension of snow and ice seasons and to an intensified production of cold arctic air masses in winter, spring and autumn. Such a process would be marked by delayed snowmelt in spring, and early snow deposition in autumn, a development not inconsistent with the fluctuations of snow covers observed from satellites in the past 12 yr^{25,26}.

We thank W. S. Broecker, W. Donn, K. Hunkins, G. Maykut, S. Manabe, J. Otterman, J. Shepherd, U. Radok, N. Untersteiner, S. Schneider, W. Wiscombe and S. Warren for comments. This study was funded by NSF grant ATM77-28522.

BHASKAR CHOUDHURY

Computer Sciences Corporation,
Silver Spring, Maryland 20910

GEORGE KUKLA

Lamont-Doherty Geological Observatory
of Columbia University,
Palisades, New York 10964

Received 2 April; accepted 14 June 1979.

1. Revelle, R. & Suess, H. *Tellus* **9**, 18–27 (1957).
2. Keeling, C. D. *et al.* *Tellus* **28**, 538–551 (1976).
3. Oeschger, H., Siegenthaler, U., Schotterer, U., & Gugelman, A. *Tellus* **27**, 168–192 (1975).
4. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **32**, 3–15 (1975).
5. Sellers, W. D. *J. appl. Meteor.* **13**, 831–833 (1974).
6. Broecker, W. S. *Science* **189**, 460–463 (1975).
7. Kellogg, W. W. *Tech. Notes Wild Met. Org.* No. 156 (1977).
8. Untersteiner, N. *Arch. Met. Geophys. Bioklim.*, A12, 151–182 (1961).
9. Bilello, M. A., Bates, R. E. & Riley, J. *NASA Tech. Rep.* **230**, 33 p. (1970).
10. Ambach, W. *Untersuchungen des Energiehaushaltes und des freien Wassergehaltes beim Abbau der winterlichen Schneedecke* (1965).
11. Lyzenga, D. R. *Icarus* **19**, 240–243 (1973).
12. Hansen, J. E. & Travis, L. D. *Space Sci. Rev.* **16**, 527–610 (1974).
13. Choudhury, B. J. & Chang, A. T. C. *NASA, Techn. Mem. TM 79639* (1978).
14. O'Brien, H. S. & Munis, R. H. *NASA SP-391*, (1975).
15. Joseph, J. H., Wiscombe, W. J. & Weinman, J. A. *J. Atmos. Sci.* **33**, 2452–2459 (1976).
16. Ivanoff, A. in *Proc. NATO Advanced Study Institute—Modeling and Prediction of the Upper Layers of the Ocean* (1975), (ed. Kraus, E.B.) 47–71 (Pergamon, New York, 1977).
17. Pfund, A. H. *J. Opt. Soc. Am.* **29**, 56–58 (1939).
18. Keeling, C. D. & Bacastow, R. B. in *Energy and Climate* 72–95 (National Academy of Sciences, Washington, D.C. (1977)).
19. McLatchie, R. A. *et al.* *USAF Cambridge Research Laboratory Tech. Rep.* 73 00 96 (1973).
20. Bo-Leckner, R. *Solar Energy* **20**, 143–150 (1978).
21. Portman, D. J. & Ryznar, E. *US Army Snow Ice and Permafrost Research Establishment Res. Rep.* **74** (1961).
22. Chernigovskii, N. T. in *Soviet Data on the Arctic Heat Budget and its Climatic Influence* (eds Fletcher, J. O., Keller, B. & Olenicoff, S. M.) 151–173 (Rand Corporation Memorandum RM-5003-PR, 1971).
23. Kojima, K., Kobayashi, D., Kobayashi, S., Naruse, R. & Ishikawa, N. *Low Temperature Science A28* (Data Rep., 1970).
24. Weaver, R. L., Barry, R. G. & Jacobs, J. D. *Proc. 3rd int. Conf. on Port and Ocean Engineering under Arctic Conditions* **1**, 455–467 (1975).
25. Kukla, G. J. *et al.* *Nature* **270**, 573–580 (1977).
26. Kukla, G. & Gavin, J. *Proc. Riederalp Workshop* (1978), IAHS-AISH Publ. no. 126, 249–258 (1979).
27. Ackerman, T. P. *Tellus* **31**, 115–123 (1979).

An apparent periodicity in an index of volcanic activity

A STOCHASTIC SIMULATION of Lamb's dust veil index^{1,2} (DVI) is used here to show that statistically significant peaks in the variance spectrum of geophysical or climatic data are not necessarily evidence of regular, or even quasi-regular, oscillations. The DVI is an annual measure of the amount of volcanic dust in the stratosphere, based on a variety of information, for the period 1500 to the present. Kelly³ has noted the significant correlation of the DVI with various aspects of climatic change in the North Atlantic–European sector on time scales of 50 yr and more, and has shown that the variance spectrum of the DVI contains a statistically significant peak at 7–8 yr. This peak is also present in the spectra of certain climatic and atmospheric circulation parameters. Kelly³ has suggested that this apparent periodicity in stratospheric dust content is more likely to be due to the chance occurrence of volcanic eruptions at 7–8 yr intervals during the limited period of record, rather than to a forced, regular cycle in volcanic activity. We now propose an alternate explanation: that the formulation of the DVI produces the spectral peak.

In his formulation of the DVI, Lamb assumes that, unless there is evidence to the contrary, after an explosive volcanic eruption the dust remains in the stratosphere in sufficient quantities to affect the radiation received at the Earth's surface, and hence climate, for four years, and that the dust amount decays linearly over that time. This assessment is largely based on the observed radiation deficits following the major eruptions of the past 100 yr and on the limited number of direct observations of the stratospheric dust load made in recent decades. Most studies of stratospheric dust loading due to volcanic activity assume an exponential, not linear, decay of the original dust injection with a relaxation time of about one year⁴. The linear decay approximates the period of most rapid fall-off of the exponential decay, integrated in yearly steps. Our results are not, however, valid for the case of an exponential decay, and the applicability of the general statistical point made here to the question of associations between volcanic activity and climate will be discussed elsewhere. The characteristic signature of a single eruption, as it appears in the DVI, is a decaying ramp function. The DVI record has an irregularity spaced saw-toothed and at times cyclical appearance due to the superposition of many such events (Fig. 1a).

Our hypothesis is that this formulation alone will produce a peak in the variance spectrum of the DVI at about 8 yr: twice the length of the ramp function. A stochastic simulation technique was used to test this hypothesis. Two parameters were allowed

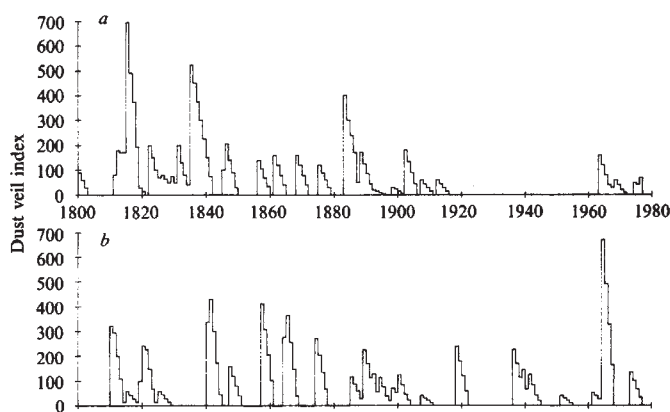


Fig. 1 *a*, Lamb's DVI: 1800–1976^{1,2}. The isolated eruptions about 1860 show the ramp formulation clearly. *b*, Stochastic simulation of the DVI, with an eruption probability of 0.3 and ramp length of 4 yr.

to vary in the stochastic simulations: the probability of an eruption within a particular year, and the ramp length. Simulated DVI records for a range of values of the two parameters were produced by a computer program using a random number generating routine. If the random number (in the range 0–1.0) assigned to a particular year was less than or equal to the eruption probability, an 'eruption' took place with a stratospheric dust injection given by a second exponentially-distributed random number. This distribution approximates to that of the actual DVI. The simulated dust-load was then linearly decayed over a period of years equal to the ramp length. A simulated DVI record is shown in Fig. 1*b*. In this case, the two variable parameters (the eruption probability and ramp length) have been set equal to those of the actual DVI. As it is a stochastic simulation the timing of the individual eruptions cannot be expected to be the same as in the actual DVI. In a series of experiments, the two parameters were varied through a realistic range of values. For each parameter pair, 1,000 records of length 117 yr were generated, the variance spectrum of each record was calculated, and a count was made of the number of peaks that were significant at the 5% level in each frequency band. The length of the simulated records used in the spectral analysis was 117 yr to facilitate comparison with the spectral analysis of the DVI given by Kelly³. The spectra were calculated using a fast Fourier algorithm, the spectral estimates were smoothed, and a white or red noise null continuum was assumed in a significance testing depending on whether or not the lag one autocorrelation coefficient departed significantly from zero.

Varying eruption probability over a wide range (0.1–0.9) produced little change in the resulting histograms (Fig. 2). This implies that a degree of superposition of the ramp functions does not obscure the spectral signal with a ramp of a few years' length. The following discussion is therefore restricted to the case of an eruption probability of 0.3 corresponding to that of the actual DVI. While this value is representative of the most reliable portion of the DVI record from 1800 to date, the eruption probability does in fact vary widely on the decadal and century-to-century time scale. The resulting histograms for ramp lengths of 4 and 6 yr are shown in Fig. 3. Clearly, as hypothesised, a spectral peak occurs at a period twice the length of the ramp function. Note that the odd harmonics of twice the length of the ramp function also appear in the histograms. As a control, we show the histogram resulting from an analysis of a set of simulations with zero ramp length (Fig. 3*c*). In this case, the total dust load was assigned to the year of the eruption and none to subsequent years. The histogram is smooth, as would be expected from a large sample of white noise spectra. This confirms that it is the ramp function alone which gives rise to the spectral signals seen in Figs 3*a* and 3*b*.

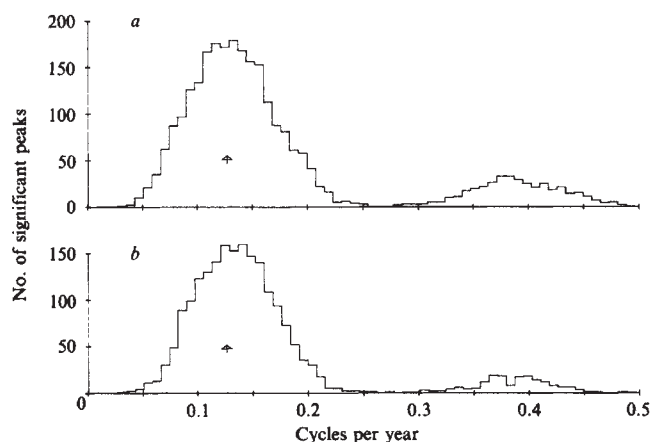


Fig. 2 Count of number of significant peaks per frequency band in the spectra of 1,000 simulations with ramp length of 4 yr and eruption probabilities of: *a*, 0.9; *b*, 0.1. Arrows indicate hypothesised location of spectral peaks.

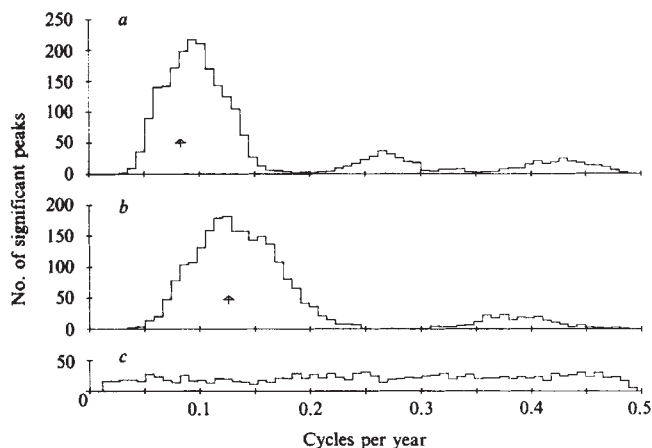


Fig. 3 Count of number of significant peaks per frequency band in the spectra of 1,000 simulations with an eruption probability of 0.3 and ramp lengths of: *a*, 6 yr; *b*, 4 yr; *c*, 0 yr. Arrows indicate hypothesised location of spectral peaks.

Two conclusions can be drawn from this analysis. First, the 7–8 yr peak in the variance spectrum of the DVI is not necessarily evidence of a regular, predictable cycle in volcanic activity. It arises from the formulation of the DVI and indicates the existence of ramp functions of length 4 yr in the data. A 7–8 yr cycle will not necessarily appear in the DVI when filtered, as the DVI is a summation of such ramp functions with random phase. This does not, however, detract from the value of the 7–8 yr spectral peak as a statistical characteristic of the DVI. It is the interpretation of this signal that is under discussion, not its reality, and it can be validly used for comparison with climatic data. The question of whether or not the linear decay is a realistic approximation to the fall-out of stratospheric dust and/or the climatic response to volcanic dust-veils, does, however, warrant further investigation. Second, we note that if there is a mechanism within the climate system which has the decay characteristics of a ramp function, for example, ocean–atmosphere interaction, then this could give rise to a spectral peak which may be misinterpreted as evidence of a regular climatic cycle.

We thank Ms B. M. Gray, Dr T. M. L. Wigley and Dr G. J. Janacek for helpful discussions, and Dr D. H. Tarling for his suggestion which prompted this analysis. This research was supported by the Office of Naval Research, grant N00014-77-G-0074.

A. CHANCE*
P. M. KELLY

*Climatic Research Unit,
School of Environmental Sciences,
University of East Anglia, Norwich, UK*

Received 9 April; accepted 10 July 1979.

* Present address: School of Computing Studies and Accountancy, University of East Anglia, Norwich, UK.

1. Lamb, H. H. *Phil. Trans. R. Soc. A* **226**, 425–533 (1970).
2. Lamb, H. H. *Clim. Mon.* **6**, 57–67 (1978).
3. Kelly, P. M. *Nature* **268**, 616–617 (1977).
4. Oliver, R. C. *J. Appl. Met.* **15**, 933–950 (1976).

The Tweeddale Granite— a newly discovered batholith in the Southern Uplands

ABOUT 2,500 gravity stations have been established in South-east Scotland by members of the Department of Geophysics, University of Edinburgh (Fig. 1). We report here a negative long-wavelength gravity anomaly over the Southern Uplands, running SW to NE parallel to the Caledonian trend. The gravity anomaly is interpreted as being due to an underlying granite

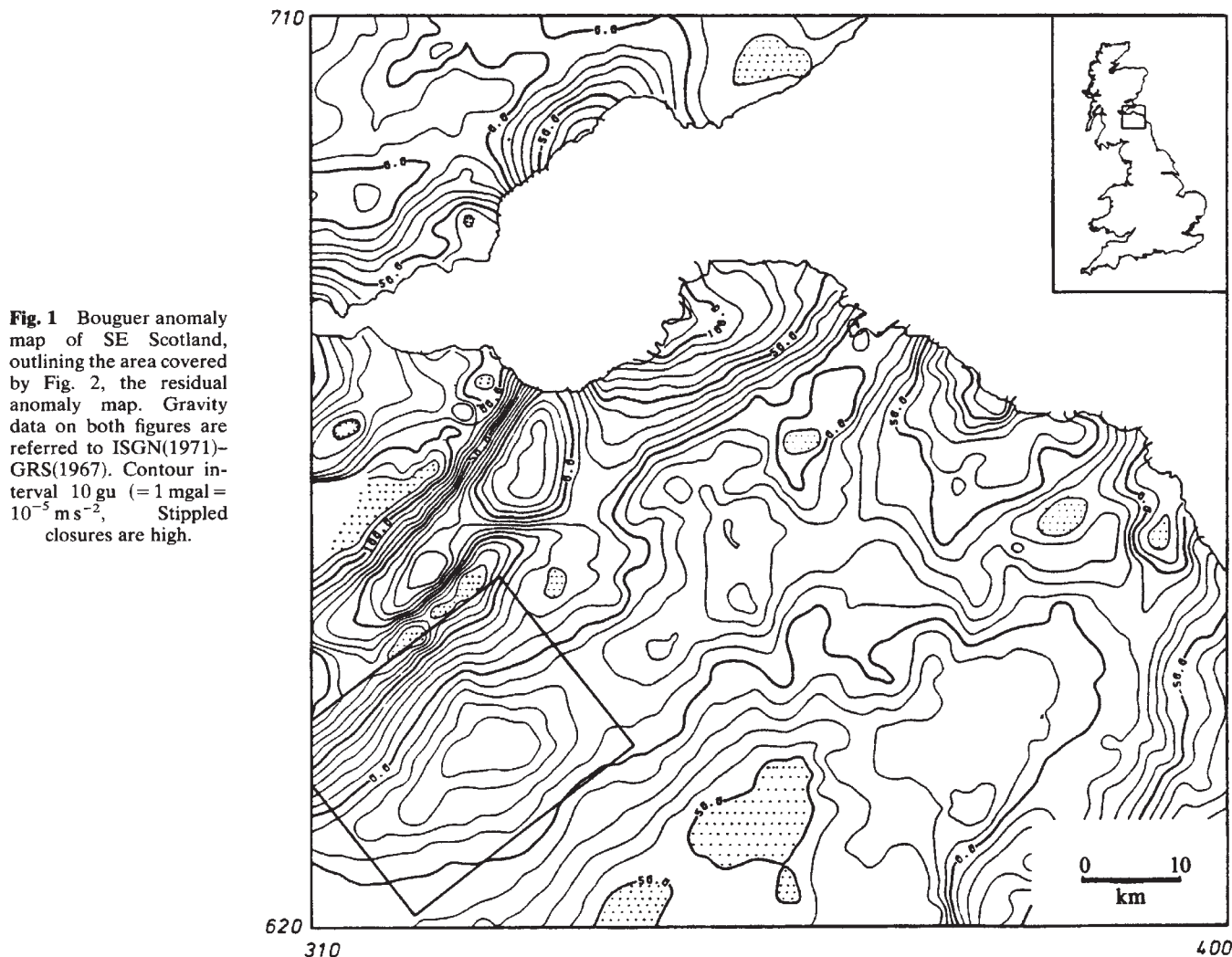


Fig. 1 Bouguer anomaly map of SE Scotland, outlining the area covered by Fig. 2, the residual anomaly map. Gravity data on both figures are referred to ISGN(1971)–GRS(1967). Contour interval 10 gu ($= 1 \text{ mgal} = 10^{-5} \text{ m s}^{-2}$). Stippled closures are high.

batholith, modelled in detail in the district of Tweeddale, where our station density averages 0.5 km^{-2} . Elsewhere in the Southern Uplands, the average coverage is 0.3 km^{-2} .

In Tweeddale the surface geological exposure to the south-east of the Southern Uplands Fault is relatively uniform. The SW–NE trending belt of Ordovician shales and graywackes is followed to the south by similar Silurian sediments as far as the Northumberland–Solway basin. Within the Lower Palaeozoic there are occasional small lenses of black shale, sometimes in the Northern Belt with radiolarian chert and pillow lavas. The whole sequence is interpreted as slices of oceanic sediment emplaced against a continental margin during the closure of the Iapetus Ocean¹. Most of the small igneous features occur as felsite and porphyrite veins, with some relatively larger exposures, notably the porphyrites at Priesthope Hill (NT3540) and at Juniper Craigs (NT2536). Although most are now only evident as boulders in the drift, some small exposures of granite have been reported² and are shown on the Geological Survey map.

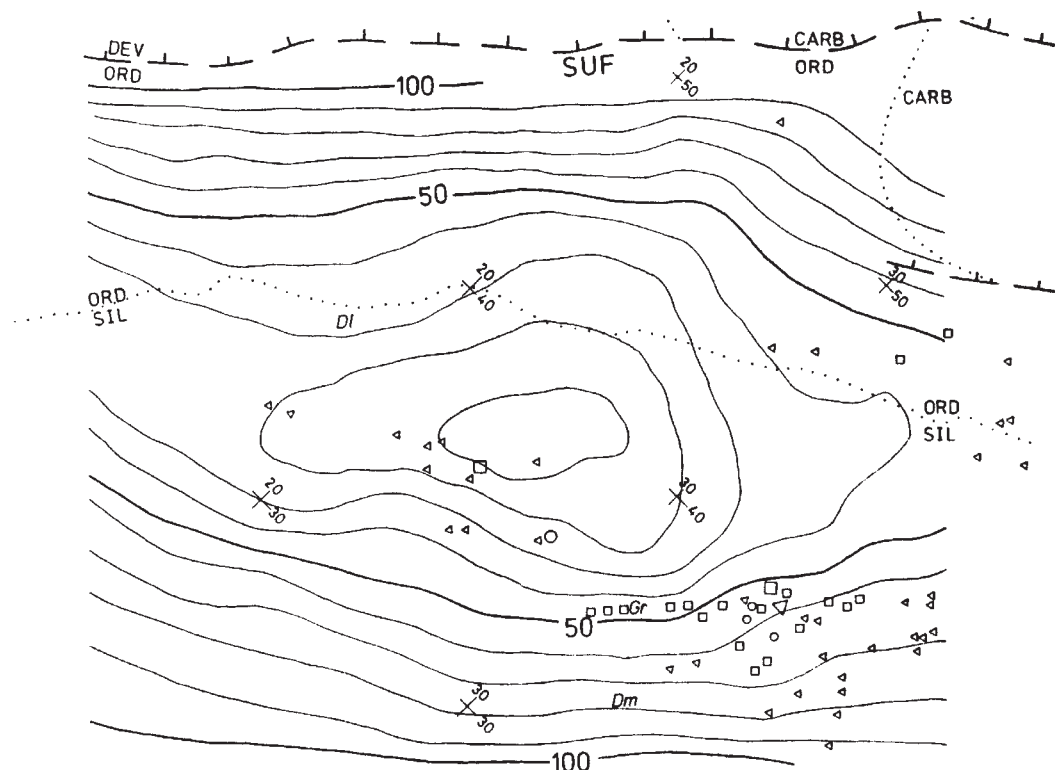
Because of the uniformity of the surface geology, the regional Bouguer anomaly field was estimated directly from the LISPB model^{3,4}, using the Nafe–Drake curve⁵ to translate seismic velocity into density. The LISPB model was assumed to be two-dimensional, striking parallel to the Caledonian trend. The resulting residual Bouguer anomaly map is shown in Fig. 2, superimposed on a geological sketch map.

Our laboratory measurements on hand samples of Lower Palaeozoic sediments from many localities in the Southern Uplands fail to show sufficient systematic variation in density to cause the gravity low. Moreover, a two-dimensional extension of Nettleton's method⁶, fitting a second degree surface separately to each 10 km National Grid square, also gave only a

small density variation: over nine squares, the computed Bouguer density was $2,711 \pm 26 \text{ kg m}^{-3}$, compared with an average from hand specimens of $2,708 \pm 21 \text{ kg m}^{-3}$. The latter was calculated from 46 samples collected from eight locations. If we attempt to model the gravity low by a variation in sediment thickness, quite excessive values ($>20 \text{ km}$) are required underlying the centre of the negative anomaly for any reasonable basement density; but it is here that the shortest wavelength features of the anomaly field occur, rather than at the margins. Evidence supporting a granite batholith is discussed below.

The centre of the main gravity low in Fig. 2 is near the outcrop of the numerous minor acid intrusives mapped in the Tweeddale area by the Geological Survey. This, in itself, suggests that the granites may be part of a more significant structure than is indicated in the field. To the SW along the Caledonian trend there are major granite exposures in South-west Scotland and Ireland, where detailed studies have already been carried out^{7–11}. Recent regional interpretation of Bouguer anomalies in South-west Scotland¹¹ suggests that the granite plutons are connected at depth as a single massive batholith extending along the Caledonian trend. The existence of a granite batholith, both in South-west Scotland and along strike in South-east Scotland, is consistent with plate tectonic models for the evolution of the British Caledonides^{12,13}. The Tweeddale area has been persistently elevated, at least since Upper Silurian times when it formed part of an emergent land mass called Cockburnland^{1,14}. The anticipated mineralisation is only sporadic in the Southern Uplands and is not strongly developed in Tweeddale: historically, lead ore was worked at a few localities near the Tweed Valley^{15,16}, most extensively at the abandoned Grieston Mine (NT3136) but also at Dalwick (NT1734) and Dumbetha

Fig. 2 Residual Bouguer anomaly map of the Tweeddale area, superimposed on a geological sketch map. Contour interval 10 gu. Intersections show the corners of 10 km National Grid squares. ORD, Ordovician; SIL, Silurian; DEV, Devonian; CARB, Carboniferous; SUF, Southern Uplands Fault; DI, Dalwick; Gr, Gries-ton; Dm, Dumbetha; Δ , felsite; $\square$, porphyrite; $\circ$, granite.



(NT3433). Geochemical analyses of organic material in the Moffat Shales¹⁷ do indicate sharply increased thermal alteration towards the north, where we model the granite intrusion.

Residual gravity anomalies, interpolated onto a regular grid, were inverted by a three-dimensional modelling program, which adjusts the topography on a single interface of contrasting density. A unique model inversion exists if the density contrast and the depth to the interface at one point are known and the depth is single-valued everywhere.

A density contrast of -70 kg m^{-3} between Ordovician and Silurian sediments and the granitic rock was assigned ($2,720$ – $2,650 \text{ kg m}^{-3}$). Published densities of exposed granitic rocks in South-east Scotland² vary between $2,630$ and $2,800 \text{ kg m}^{-3}$ and do little to constrain the interpretation. From a few massive boulders on Kailzie Hill (NT2836) and Kernie Law (NT3439), we found a saturated density of $2,628$ ($2,644 \pm 9 \text{ kg m}^{-3}$) (11 samples), with $2,663$ ($2,674 \pm 27 \text{ kg m}^{-3}$) (42 samples) for nearby porphyrites and felsites (grain densities in brackets). In South-west Scotland, the Fleet granite has been modelled¹⁰ with a density of $2,630 \text{ kg m}^{-3}$, while the Griffler modelirite was interpreted⁸ with a gradational density between $2,640$ and $2,710 \text{ kg m}^{-3}$.

The depth-scale is more uncertain: a three-point depth rule¹⁸ suggested that the interface is less than 2.6 km deep near the centre of the anomaly, while the maximum amplitude/maximum gradient ratio put the centre of mass at $< 6.5 \text{ km}$. Figure 3 shows a typical three dimensional model inversion, here with a central point fixed at 2 km depth. The shape is typical for a batholith with outward sloping sides extending to a depth of 10 – 12 km and three small bosses coming near to the surface close to the porphyrite outcrops. This model is consistent with the barycentre estimate as well as seismic and magnetotelluric models showing a boundary at about 12 km depth.

Ingham (personal communication) finds a highly resistive layer ($10^4 \Omega \text{m}$) overlying a good conductor ($10^2 \Omega \text{m}$) at a depth of 10 – 12 km , from a one-dimensional inversion of magnetotelluric data from sites at Earlyburn (NT2249), Peebles (NT2740), Yarrow (NT2924) and Borthwick Brae (NT3616). The high overlying resistivity ($10^4 \Omega \text{m}$) is indicative of a low-porosity igneous or metamorphic rock and so is also consistent with our granite hypothesis. Jacob¹⁹ interpreted seismic arrivals at the

Eskdalemuir array in terms of a 6.44 km s^{-1} refractor at a depth of 12 km , a model which is not inconsistent with the later LISP result³, where this mid-crustal refractor was only poorly constrained. Although the LISP profile actually crosses the Tweeddale area, it should not be expected to show the lateral

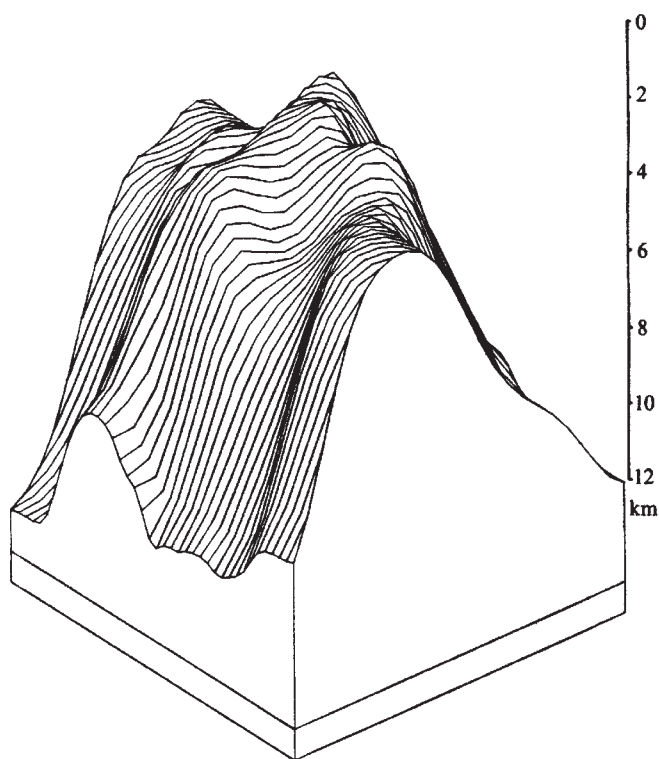


Fig. 3 Three-dimensional model of the upper surface of the Tweeddale Granite, viewed from the west and covering the same area as the residual anomaly map in Fig. 2. The computation includes infinite horizontal extensions from all edges of the model, the beginnings of which in the SW–NE direction are included in the drawing.

extent of the granite batholith: the $5.84 \pm 0.02 \text{ km s}^{-1}$ velocity observed for the supposedly Lower Palaeozoic sediments does not contrast with values found generally for granites^{20,21}, although the S-wave arrivals do give an unusually low Poisson's ratio²², consistent with quartz enrichment.

Figure 1 suggests that the negative gravity anomaly, and therefore the granite batholith, extends northeastwards from the Tweeddale area towards the Berwickshire coast, although the pattern is confused by the Devonian sedimentary basins in Lauderdale and near Oldhamstocks (NT7471) and Eyemouth (NT9464). From the aeromagnetic map²³, and more clearly from filtered versions of it^{24,25}, there is reasonable long-wavelength continuity of a weak magnetic high, flanked to the north by a corresponding low, from the Criffel-Fleet-Doon plutons in SW Scotland to the Tweeddale area and thence northeastwards towards the coast. The 'granites' of Berwickshire and E. Lothian^{2,26} show close petrological affinities to those in South-west Scotland. Only the Priestlaw (NT6562) and Lammerlaw (NT5163) exposures show localised, stocklike, positive magnetic anomalies, although all appear to be related to the same weak regional magnetic high traceable along the whole NE-SW length of the Southern Uplands. It seems reasonable to postulate that the whole Southern Uplands are underlain by the batholith, which is responsible for their longstanding and continued relatively high elevation and stability.

E. LAGIOS
R. G. HIPKIN

Department of Geophysics,
University of Edinburgh, UK

Received 4 June; accepted 16 July 1979.

- McKerrow, W. S., Leggett, J. K. & Eales, M. H. *Nature* **267**, 237-239 (1977).
- Walker, F., *Trans. Edinb. geol. Soc.* **11**, 357-365 (1924). Walker, F. *Geol. Mag.* **65**, 153-162 (1928).
- Bamford, D., Nunn, K., Prodehl, C. & Jacob, B. *J. geol. Soc. Lond.* **133**, 481-488 (1977).
- Bamford, D., Nunn, K., Prodehl, C. & Jacob, B. *Geophys. J. R. astr. Soc.* **54**, 43-60 (1978).
- Nafe, J. E. & Drake, C. L. in *The Sea* Vol. 3 (ed. Hill, M. N.) 794-815 (Interscience, New York, 1963).
- Nettleton, L. L. *Geophysics* **4**, 176-183 (1939).
- Cook, A. H. & Murphy, T. *Geophys. Mem. Dubl.* **2**(4) (1952).
- Bott, M. H. P. & Masson Smith, D. *Proc. Yorks geol. Soc.* **32**, 317-332 (1960).
- Parslow, G. R. *Scott. J. Geol.* **9**, 91-108 (1968).
- Parslow, G. R. & Randall, B. A. O. *Scott. J. Geol.* **9**, 219-231 (1973).
- El-Batroukh, S. I. thesis Univ. Glasgow (1975).
- Mitchell, A. H. G. & McKerrow, W. S. *Geol. Soc. Am. Bull.* **88**, 764-788 (1975).
- Phillips, W. E. A., Stillman, C. J. & Murphy, T. *J. geol. Soc. Lond.* **132**, 579-609 (1976).
- Walton, E. K. in *The Geology of Scotland* (ed. Craig, G. T.) 161-227 (Oliver & Boyd, Edinburgh, 1965).
- Hardy, J. *Hist. Berwicksh. Nat. Club*, **9**, 481 (1892).
- Wilson, G. V. & Fleet, J. S. *Mem. geol. Surv. spec. Rep. Miner. Resour. Gt Br.* **17**, 62-63 (1921).
- Watson, S. W. thesis Univ. St. Andrews (1976).
- Bott, M. H. P. & Smith, R. A. *Geophys. Prospect.* **6**, 1-10 (1958).
- Jacob, A. W. B., *Geophys. J. R. astr. Soc.* **18**, 189-197 (1965).
- Woollard, G. P. *Res. Rep. Ser. 62-9*, Univ. Wisconsin (1962).
- Holder, A. P. & Bott, M. H. P. *Geophys. J. R. astr. Soc.* **23**, 465-468 (1971).
- Assumpcao, M. & Bamford, D. *Geophys. J. R. astr. Soc.* **53**, 61-73 (1978).
- Bullerwell, W. *Aeromagnetic Map of Great Britain* (1970).
- Hall, D. H. & Dagley, P. *Inst. Geol. Sci. Rep.* **70/10** (1970).
- Gunn, P. J. thesis Univ. Durham (1972).
- Bennett, J. R. P. *Inst. Geol. Sci. Rep GP/AG/70/13* (1969).

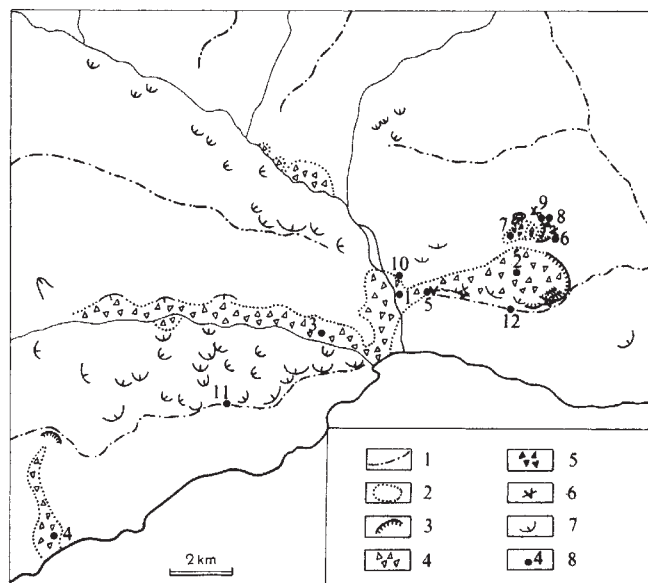


Fig. 1 Scheme of seismic dislocations and sites of lichenometric measurements in the epicentral area of the 1949 Khait earthquake. 1, Watersheds; 2, contours of landslides; 3, fracture walls of rock landslides; 4, landslides of 1949; 5, ancient landslides; 6, seismoravines in basement rock; 7, earth avalanches and slips of (mainly loess) rock; 8, sites of lichenometric measurements.

year of exposure. We used lichens of species *Lecidea lactea* Ach. growing approximately in circles, which are convenient shapes for measurements. Each sample included, wherever possible, 100 measurements of the lichen diameter on compact groups of homogeneous rock surfaces.

The investigation was in two stages: (1) the average growth rate of lichens on geologic objects of known age was determined, and (2) the age of undated objects was determined from measurements of the lichen diameter using the average growth rate values obtained.

The area of the 1949 Khait earthquake lies on the right bank of the Surkhob river in the outskirts of the village of Khait. The earthquake ($M = 7.4 \pm 0.1$) had caused a large landslide and an avalanche (length 10 km, height above 50 m) of mud and stone, as well as producing seismic ravines in hard rock²⁻⁴. To determine the average growth rate of lichens in this area, its dependence on exposure conditions and altitude (the ecology conditions) had to be estimated first. Two series of measurements were taken of boulders and debris in the upper and lower portions of a landslide^{2,3} in the epicentral area of the 1949 Khait earthquake. The diameter distribution curves of lichens inhabiting the southern and northern surfaces of the boulders and debris were practically identical. Thus, the rate of growth can be taken to be independent of exposure conditions, but is highly dependent on altitude (Table 1) and this should be included in the calculations.

Table 1 shows that a number of statistical parameters of the rate of growth have been used: the arithmetic mean of diameter values (X_{mean}), the diameter of the most frequently occurring

Table 1 Basic statistical characteristics of average yearly growth (in mm) of lichens of species *Lecidea lactea*

Statistical characteristics of growth from diameter measurements	Khait area (mm yr^{-1})		Karatag area (mm yr^{-1})
	Site 1 Altitude (m)	Site 2 Altitude (m)	Site 1 Altitude (m)
	1,600	2,000	1,700
Arithmetic mean, X_{mean}	0.54	0.47	0.38
Most frequent, X^*	0.48	0.40	0.38
Largest, C_{max}	1.20	0.94	—
Maximum, X_{max}	1.30	1.00	0.75
Yearly change of variance	0.14	0.13	0.092

Lichenometry and earthquake age determination in central Asia

THE age of rock surfaces can be determined from the size of lichens growing on them¹. We present here measurements of the species *Lecidea lactea* growing on rocks from the avalanches which occurred after the 1907 (Karatag) and the 1949 (Khait) earthquakes in Tadzhikistan, Middle Asia. We show that their forerunning strong earthquakes occurred more than 100-130 yr previously.

The method of using lichens to date rocks is based on two assumptions: first, that the yearly rate of lichen growth is constant (on average within a given area), and second, that lichens settle on a newly exposed rock surface during the first

Table 2 Lichenometric characteristics and age of various objects on the epicentral zone of the 1949 Khait earthquake

Group	Measure- ment sites	Description of object	Altitude (m)	Sample size	Statistical characteristics of lichen diameter (mm above) and age determinations (yr below)					Average age and confidence interval (yr)	Time of event (yr)
					X_{mean}	X^*	σ	X_{max}^*	X_{max}		
(1)	3	Boulders in landslide on right bank of Yagman river 2.5 km west of village of Khait	1,600	100	13.1	13.5	3.8	30.5	30.5	25 ± 1	1950 $\pm$ 55
					24	28	27	25	23		
(2)	4	Boulders in landslide on right bank of Surkhob river around village of Mirtob	1,500	100	15.5	11.5	6.4	36.5	36.5	30 ± 7	1940–50
					27	23	44	28	27		
	5	Ravine walls in western part of latitudinal watershed 0.5 km NE of village of Khait	1,650–1,750	97	15.8	14.5	4.5	31.5	32.5	29 ± 3	1940–50
					30	31	33	27	26		
	6	Renewed ravine walls in northern part of meridional watershed 5 km NE of village of Khait	3,000	16	13.9	13.5	2.3	—	15.8	27 ± 7	1940–60
					27	36	19	—	—		
	7	Ancient landslide 4 km NE of village of Khait	2,200	42	75.0	88.0	17.5	—	—	170 ± 70	1740–1880
					160	220	140	—	—		
(3)	8	Walls of ancient ravine on meridional watershed 5 km NE of village of Khait	2,800	10	99.5	75.0	24.0	132.5	132.5	190 ± 10	1780–1800
					200	190	190	—	—		
	9	Walls of another ancient ravine, same location	2,800–3,000	22	71.1	73.0	20.0	112.5	113.0	160 ± 20	1800–40
					150	180	160	—	—		
(4)	10	Ancient exposed surface near the foot of a slope down to Yarkhych river, 0.7 km NE of village of Khait	1,600	62	102.8	94.0	26.0	170.0	176.9	170 ± 30	1780–1840
					190	195	185	140	130		
	11	Rock exposures at the ridge of a range 6 km WSW of village of Khait	2,400–2,450	50	92.1	82.5	25.0	160.0	163.9	190 ± 20	1770–1810
					195	210	220	170	165		
	12	Rock exposures at a latitudinal ridge 3 km east of village of Khait	2,500	37	107.4	105.0	25.0	195.0	197.5	220 ± 30	1730–90
					230	206	200	210	200		

Table 3 Lichenometric characteristics and age of various objects on the epicentral zone of the 1907 Karatag earthquake

Group	Measure- ment sites	Description of object	Altitude (m)	Sample size	Statistical characteristics of age determinations (yr) from				Average age and confidence interval (yr)	Time of event (yr)
					X_{mean}	X^*	X_{max}^*	X_{max}		
(1)	2	Boulders in a small landslide on right bank of Karatag valley 1 km east of village of Chuyanchi	1,700	28	62	49	89	51	63 ± 22	1890–1930
					77	—	71	49		
(2)	3	Walls of a seismoravine on watershed near village of Chuyanchi	1,800	6	69	69	61	60	65 ± 6	1910–20
					32	29	28	27		
	5	Boulders in a fresh landslide on left bank of Karatag river 2.5 km NNE of northern edge of village of Karatag	1,150	63	28	24	34	21	27 ± 7	1940–60
					105	84	174	130		
	7	Boulders in an ancient landslide on left bank of Karatag river 4 km NE of northern edge of village of Karatag	1,700	73	238	236	182	173	212 ± 41	1720–1810
					203	203	149	157		
(3)	8	Boulders in an ancient landslide on right bank of Karatag river above village of Naalbek	1,600	47	203	203	149	157	178 ± 34	1770–1830
					203	203	149	157		
	9	Ancient seismoravine on left bank of Karatag river 3 km NNE of northern edge of village of Karatag	1,200–1,300	38	203	203	149	157	178 ± 34	1770–1830
					203	203	149	157		

The average age has been calculated for a reference height of 1,700 m.

objects (X^*), the mean diameter of the largest objects (X_{max}^*), the maximum diameter (X_{max}). We have also estimated the variance (mean squared deviation) (σ). We believe that such an approach greatly improves the accuracy attainable using the maximum diameter measurements alone^{1,5} and confirms the desirability of combining several characteristics of growth rate⁶.

Before lichenometry could be applied to objects of unknown

age, they had to be classified into several groups according to outward appearance and geomorphologic criteria: (1) boulders in landslides probably dating from the 1949 earthquake (from indirect indications see refs 2, 3, and outward appearance); (2) walls in ravines that were either produced or renewed by the earthquake (our unpublished data); (3) ancient ravines and a landslide that were probably produced by a large earthquake

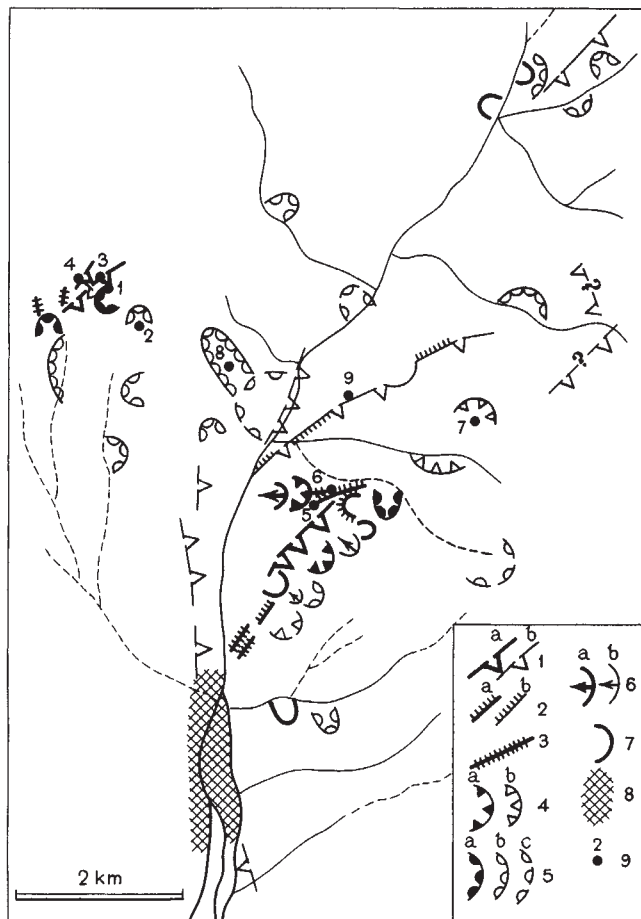


Fig. 2 Scheme of seismic dislocations and sites of lichenometric measurements in the epicentral zone of the 1907 Karatag earthquake. 1, Seismoravines; 2, terraces, escarpments; 3, cracks in loose sediments; 4, fracture walls of rock landslides; 5, landslides; 6, earthavalanches and slips of loose rock; 8, village; 9, sites of lichenometric measurement. Age by geomorphologic criteria: a, fresh; b, somewhat earlier (late Holocene); c, ancient (pre-Holocene).

before 1949 (our unpublished data); (4) hard rock surfaces that are certainly older than the 1949 or any previous earthquake, or which were renewed during it.

Two or three objects from each group have been measured. Table 2 shows that recent landslides in the Yagman river valley and a part of the Surkhob valley around the village of Mirtob were actually formed during the 1949 earthquake, the latitudinal ravines in the watershed to the east of Khait were indeed formed during the 1949 earthquake, whereas the meridional ravines are much older, and were only renewed in 1949; the ancient landslide and ravines in the upper Daraihaus river are approximately of the same age, so that they can be considered as paragenetic formations of an earlier large seismic event (or several consecutive events); the earthquake-generated surfaces are younger than the other hard rock surfaces, that is the age of the early seismic dislocations in our study is not limited by the extreme size of lichens in the present conditions.

Thus, we may conclude that the epicentral area of the 1949 Khait earthquake had been struck by a similar sized earthquake a long time previously.

The 1907 Karatag earthquake area ($M = 7.4 \pm 0.1$) faces the southern foot of the Hissar range 50 km west of the town of Dushanbe. The objects studied by lichenometry in the epicentral area of the Karatag earthquake can also be classified into a number of groups: in addition to the wall of a joint that certainly results from the Chuyanchi earthquake of 21 October 1907 (ref. 7), and which is used here to determine the average growth rate of lichens in this part of the area (see Table 1), there are the following: (1) ravine walls and landslide boulders which we

tentatively ascribe from their positions and outward appearance to the 21 October 1907 earthquake; (2) boulders in the landslide and walls in the ravine, supposedly a result of the 8 October 1907 earthquake; (3) boulders in ancient landslides and the walls of an ancient ravine, here associated with an earthquake(s) before 1907.

The results of lichenometric measurements in the Karatag area support the idea⁸ that the landslide and ravines near the village of Chuyanchi (group (1)) are associated with the Chuyanchi earthquake of 21 October 1907. The landslides and ravines on the left bank (group (2)), which we had supposed to originate from the Karatag earthquake of 8 October 1907 have turned out to be much younger (29 ± 3 yr and 27 ± 7 yr from the age of the lichens). We consider, however, that this age only reflects a renewal of the surface rather than formation of the objects themselves.

The group (3) formations ascribed to an earlier earthquake some hundreds of years before 1907 (refs 8, 9), have an age of 200 yr according to lichenometry. A landslide on the left bank first investigated here (measurement site 7 in Table 3) can be tentatively associated with the large 1863 earthquake west of the Hissar range from some age determinations from the lichens on the boulders (123 ± 43 yr).

The two largest earthquakes that we know to have occurred in the Hissar-Kokshaal fault zone in Tajikistan cannot be considered to be random events as regards their place and time of occurrence. The investigation area has been struck with large ($6 < M \leq 7.5$) earthquakes during the past few hundred years with an interval of recurrence of at least 100–130 yr, but more likely 200–250 or more years.

A. A. NIKONOV
T. YU. SHEBALINA

*Institute of Physics of the Earth,
B. Gruzinskaja 10,
810123 Moscow, USSR*

Received 4 January; accepted 27 June 1979.

1. Beschel, R. E. *Arctic* 7 (1957).
2. Gubin, I. E. *Certain Regularities in the Seismic Activities of Tajikistan* (Academy Science USSR, Moscow, 1960).
3. Leonov, N. N. *Izv. Acad. Sci. USSR, Geophys. Ser.* (1960).
4. Solonenko, V. P. in *Geologic Regularities in and Mutual Relations of Land-slides, Avalanches, and Mountain Mud Flows*, Part II (MGU, 1976).
5. Webber, R. I. & Andrews, I. T. *Arctic and Alpine Res.* (1973).
6. Benedict, B. J. *Glaciol.* 7, 77–78 (1967).
7. Korolkov, B. *Izv. Turkm. sect. Russian geogr. Soc.* 9 (1913).
8. *Recent Seismic Dislocations and their Bearing on the Problem of Seismic Microzoning* (MGU, 1977).
9. Nikonov, A. A. *Holocene and Recent Crustal Movements* (Nauka, Moscow, 1977).

Do circadian oscillators ever stop in constant light?

A TWO-DIMENSIONAL limit cycle model (Fig. 1) predicts the behaviour in darkness of both the *Drosophila pseudoobscura* circadian eclosion rhythm^{1,2} and the mosquito *Culex pipiens quinquefasciatus* Say (*C. p. fatigans* Wied.) flight activity rhythm^{3,4}. This model has been extended to explain features of rhythms in very dim light, but in constant 'bright' light it has been suggested that the dynamics are fundamentally different, with the system having a globally stable equilibrium point and hence no stable oscillation¹. We propose here a simplifying modification: for each light intensity the dynamics are topologically equivalent, and thus for each there is a limit cycle, its position and amplitude being continuous functions of intensity. This interpretation avoids the awkward discontinuity between 'light' and 'dark' dynamics and emphasises the possibility of time-keeping at all light intensities. Furthermore, it provides a simple explanation for novel features of the *Culex* rhythm described here and previously³, and allows us to treat some

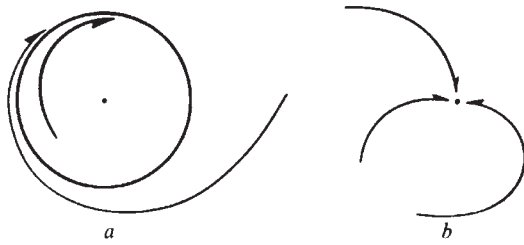


Fig. 1 *a*, Outline of the two-dimensional limit cycle model. The two variables are left unspecified, the oscillator is merely assumed to be attracted in darkness to a closed cycle, here represented by the bold circle. This cycle, the limit cycle, is assumed to be stable. If the oscillator is displaced from the cycle (typically by exposure to light), it will return as indicated by the two sample trajectories. There must be at least one point, the critical point, within the region bounded by the cycle, from which this return is indeterminate. *b*, A globally stable equilibrium point. All trajectories are attracted to a single point and thus such a system will always eventually reach a fixed non-oscillatory state.

important differences between *Culex* and *Drosophila* as merely quantitative. Finally, it suggests a new general strategy for the analysis of circadian rhythms and allows extensions of established theory that explain previously paradoxical results obtained for the two species (E.L.P. in preparation).

Both the *Drosophila* and *Culex* rhythms persist in continuous darkness, following, according to most models, a limit cycle in the two-dimensional state space. Phase is assigned to all limit cycle states relative to an arbitrary reference point. States out of the cycle can only be defined operationally: for example, that state realised after a phase 0 population is given a 1 s, 1 lx perturbation. However, states can be compared with the following protocol: transfer samples to darkness and let the original states be equivalent if the settled rhythms in darkness are in phase. The state space is thereby partitioned into equivalence classes or isochrons² (Fig. 2). Each isochron must include a unique limit cycle state and is labelled accordingly (isochron θ contains the state with phase θ). In practice, a state's isochron is determined by an assay experiment entailing transfer to darkness and comparison with a sample having known phase. In fact, not quite all states belong to an isochron, from some the return of the rhythm is unpredictable or not apparent. Such states comprise the phaseless set, which must include at least one point in the region bounded by the limit cycle.

The simplest possible topology is illustrated by

$$\begin{aligned}\dot{x} &= y + \varepsilon x(a^2 - x^2 - y^2) \\ \dot{y} &= -x + \varepsilon y(a^2 - x^2 - y^2)\end{aligned}\quad (1)$$

which has a limit cycle of radius a , centred on the phaseless critical point at the origin and straight, equally spaced radial isochrons. Each trajectory moves with constant angular velocity about the origin, independently of a and ε .

Our working hypothesis is that in darkness the circadian dynamical system has qualitatively the same form as equation (1). Note that in such cases, the isochrons divide state space much as longitudinal lines divide the Northern Hemisphere (Fig. 2), both systems yield incomplete yet useful information about location. The release-assay experiment uses this information to map a trajectory's route through state space: a population is synchronised at a chosen state, the release point, then released to free-run in constant test conditions—light of fixed intensity. The population follows the trajectory dictated by the dynamics at that light intensity. By periodically removing an aliquot and assaying the phase in constant darkness, the population's progress across the isochron map can be monitored. A plot of isochron against time gives an assay curve.

While this formal presentation and interpretation of release-assay experiments is new, the experimental design is not. Pittendrigh's *Drosophila* experiment⁵ is one archetype (Fig. 2*a*). The dynamics in bright light apparently force trajectories to a

point on one isochron, which, for this reason, we have defined to be isochron 0.

With these two concepts—a limit cycle in darkness and a globally stable equilibrium point in light—one can explain most light pulse phase resetting phenomena in *Drosophila*² and *Culex*. Yet, because the trajectory patterns in the two cases are topologically different (one with a limit cycle, one without), somewhere on the continuum of light intensity a change of form must occur. We prefer to eliminate the discontinuity, proposing instead that the trajectory patterns for all light intensities are topologically equivalent and, in particular, that a limit cycle always exists. For example,

$$\begin{aligned}\dot{x} &= (y - b) + \varepsilon x(a^2 - x^2 - (y - b)^2) \\ \dot{y} &= -x + \varepsilon(y - b)(a^2 - x^2 - (y - b)^2)\end{aligned}\quad (2)$$

where $a = \exp(-kI)$ and $b = y_\infty(1 - \exp(-kI))$ is a family of systems parametrically defined by light intensity, I . Each member of equation (2) is just a coordinate shift of a corresponding member of equation (1). Hence for each light intensity there is a limit cycle, radius a , centred on a single critical point $(0, b)$, the angular velocity of every trajectory around the critical point for that intensity is constant. As intensity tends to infinity, the corresponding system tends

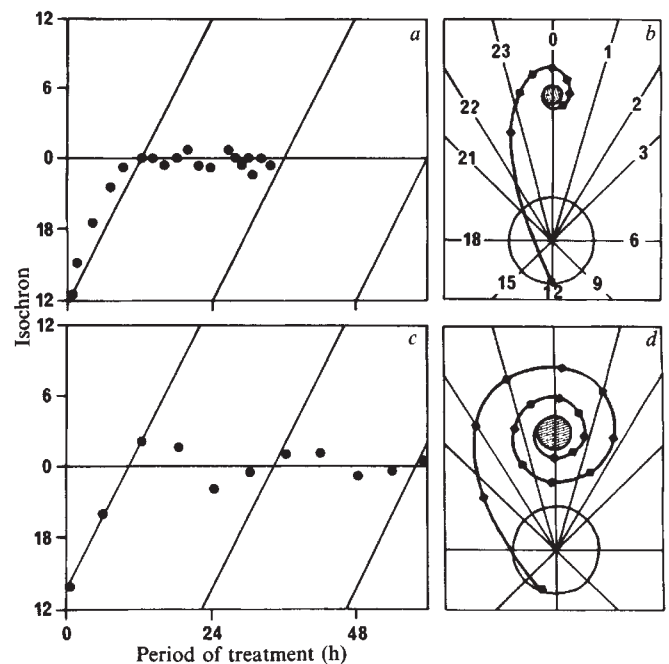


Fig. 2 *a, c*, Assay curves. The set of diagonals comprise the assay curve which would necessarily be obtained if the treatment condition were darkness and correspond to a population continuously following the dark limit cycle. *b, d*, Corresponding route maps. The route map is an interpretation of the associated assay curve. Each includes the dark limit cycle (unshaded circle), radial isochrons and the limit cycle assumed to exist in the given test conditions (shaded circle). Superimposed is the trajectory of the sample, inferred from the assay curve, assuming the dynamics are approximately those of equation (2). The predicted position of the sample every 3 h is indicated on the trajectory. *a*, *Drosophila pseudoobscura* assay curve. Data are redrawn from Fig. 7 of ref. 5. The release point was 'light on' in LD 12:12 (12 h light, 12 h dark), the test condition was 1,080 lx. Isochron 0 to which the assay curve seems to tend corresponds to Pittendrigh's ct: 12. *b*, *Drosophila pseudoobscura* route map. The proposed limit cycle is so small that any oscillation would be masked by the large experimental uncertainty. *c*, *Culex pipiens quinquefasciatus* assay curve redrawn from Fig. 2 of ref. 3. The release point was 'light on' in LD 12:12, the test condition was 50 lx. Each point is a mean of 24 mosquitoes with s.e. < 15 min. (The error bars in ref. 3 Fig. 2 are standard deviations.) *d*, *Culex pipiens quinquefasciatus* route map. The assay curve moves rather more quickly to its first maximum (12 h) than predicted, assuming constant angular motion around the 50 lx critical point.

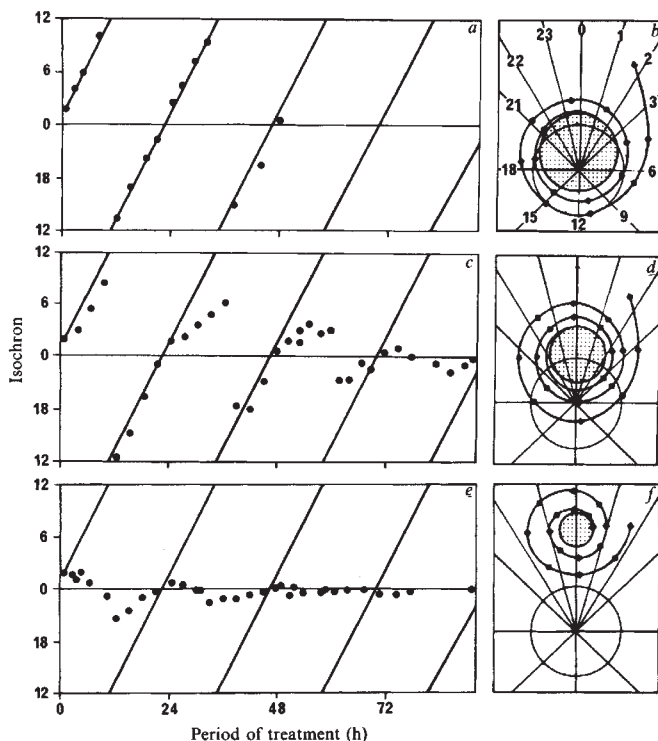


Fig. 3 The general organisation is as in Fig. 2. 'Light off' in LD 12:12 (180 lx) was chosen as the release point. (We have noted no difference between 50 and 180 lx.) Activity was monitored acoustically⁴ and each point is the mean for 24 mosquitoes. *a, b*, 0.05 lx. The population follows the dark assay curve for the period examined. We infer that the 0.05 lx limit cycle is rather close in amplitude and position to the dark limit cycle. *c, d*, 0.5 lx. During the first 24 h the assay curve is near the dark assay curve. Correspondingly, the first cycle of the route map trajectory curls around the singularity cutting all isochrons. Over the second 24 h the assay curve jumps from isochron 6 to isochron 18, a result predicted if the trajectory cuts across the singularity at this time. Subsequent turns of spiral exclude the singularity, thereby failing to cut most isochrons. This is reflected by the reduced excursion of the assay curve in later cycles. *e, f*, 8 lx. The data points of Fig. 2c fall almost perfectly on those of *e* (after a 12-h time shift) indicating that the dynamic systems corresponding to 8 and 50 lx probably differ very little. The limit cycle, if it exists at this intensity, is too small to detect above experimental error, although the transient spiral is readily apparent.

smoothly toward the limiting case—the system with a stable focus at $(0, y_{\infty})$.

This scheme is compatible with the *Drosophila* result, assuming that the amplitude of the limit cycle is sufficiently reduced for the oscillation to be experimentally irresolvable (Fig. 2b). However, it was the comparable experiment for *Culex*³ (Fig. 2c) which yielded tangible evidence for the model. A spiral trajectory asymptotically approaching a limit cycle (Fig. 2d) is the only plausible explanation for the form of the assay curve, and the persistence of a circadian activity rhythm in the test conditions³.

We propose that the organisation of these two dipteran circadian systems is qualitatively similar, discrepancies reflecting differences in treatment intensity and light sensitivity in the two cases. In support of this, note that a circadian eclosion rhythm continues in *Drosophila* in very dim light (0.05 lx white light⁶ or 0.01 erg cm⁻² blue light²), although an assay curve or its equivalent has never been presented.

The set of dim light assay curves for *Culex* suggests the smooth transitional sequence between darkness and light demanded by the model (Fig. 3). The curve for 0.5 lx (Fig. 3c) is of critical significance. There is, between 36 and 39 h, a phase jump of about 12 h, a feature compatible with the model only if the trajectory passes near the singularity of the isochron map

during that time. Winfree² reports that for *Drosophila* the region around the singularity is a phaseless zone from which a return to rhythmicity is not observed in the dark. Results suggesting a similar phenomenon for *Culex* were seen in the 36 and 39 h assays. Depressed arrhythmic activity was recorded in these two samples, assumed to be lingering near the singularity. Although to some extent the result of population incoherence, the effect is obvious in individual activity records. This damping effect, like that reported for *Drosophila*, is reversed by a period of bright light, as the model requires.

We first proposed that the systems had straight isochrons merely to make diagrams easy to draw and predictions correspondingly clear. Even so, in addition to the phenomena considered here, other features including phase resetting, steps up and down in light intensity and entrainment can be explained without seriously weakening this simplifying assumption. Our philosophy is thus to maintain this general framework until an obvious inconsistency appears.

The existence in *Culex* of activity rhythms in light suggests an obvious test of the model's consistency. Choosing any intensity one can select a phase marker, partition the state space into a new set of isochrons corresponding to that intensity, and conduct release-assay experiments using this new isochron map as a reference frame. Test conditions can be any light intensity including darkness. The form of these new assay curves can be predicted from that of the old curves assuming constant angular motion. Such experiments are in progress.

We have considered here the dynamics of a single oscillator, although it is almost certain that circadian rhythms reflect the interaction between coupled oscillators. Nevertheless, as Pavlidis⁷ argues, preserving this duality in the description of circadian rhythms is useful and leads to no important contradiction.

We thank Dr B. C. Goodwin for helpful suggestions and criticisms of the manuscript. The work was supported by grants from the MRC.

ERIC L. PETERSON
M. D. R. JONES

School of Biological Sciences,
University of Sussex,
Falmer, Brighton, Sussex, UK

Received 24 May; accepted 13 July 1979.

1. Pavlidis, T. *Lectures on Mathematics in the Life Sciences* Vol. 1 (ed. Gerstenhaber, M.) 88–112 (Am. Math. Soc., Providence, Rhode Island, 1968).
2. Winfree, A. T. *J. theor. Biol.* **28**, 327–374 (1970), *J. comp. Physiol.* **85**, 105–140 (1973), *Science* **183**, 970–972 (1974).
3. Jones, M. D. R. *Nature* **261**, 491–492 (1976).
4. Peterson, E. L. *Behaviour* (in the press).
5. Pittendrigh, C. S. *Z. Pflanzenphysiol.* **54**, 275–307 (1966).
6. Pittendrigh, C. S. *The Molecular Basis of Circadian Rhythms* (eds Hastings, J. W. & Schweiger, H.-G.) 85–108 (Dahlem Konferenzen, Berlin, 1976).
7. Pavlidis, T. *Bull. math. Biol.* **40**, 625–635 (1978), 675–692 (1978).

Relative 'trophic' influences of excitatory and inhibitory innervation of locust skeletal muscle fibres

DENERVATION of locust skeletal muscle, where L-glutamate is thought to be the excitatory neurotransmitter¹, induces a pattern of physiological and pharmacological changes similar to that following denervation of vertebrate skeletal muscle^{2,3}. The degeneration of nerve terminals on locust muscle which follows axotomy⁴ is accompanied by a significant increase in the sensitivity of the extrajunctional membrane to L-glutamate^{5–8}, as well as a fall in resting membrane potential^{8,9} and input conductance⁸ of the denervated fibres. Locust skeletal muscle fibres are multiterminal and sometimes polyneuronally innervated. In addition to excitatory inputs some fibres receive endings from inhibitory axons in which transmission is probably mediated by γ -aminobutyric acid (GABA)¹⁰. This dual nervous control at the

peripheral level allows examination of the separate 'trophic' influences of excitatory and inhibitory inputs to a muscle fibre. In this report we demonstrate that in at least one locust muscle, the metathoracic retractor unguis (RU) of *Schistocerca gregaria* which contains fibres receiving both excitatory and inhibitory inputs, denervation-induced changes can occur in the presence of functionally intact inhibitory innervation.

The femoral portion of the RU muscle consists of two bundles of long (1 cm) fibres, called 'red' and 'white', usually containing eight and nine fibres, respectively^{4,11}. Each bundle of fibres is innervated by one of a pair of large ('fast') excitatory motor axons which run in a branch (B_2) (ref. 4) of metathoracic nerve 5 (Fig. 1). Previous studies have suggested that the innervation of the RU is restricted to excitatory neurones; however, antidromic stimulation of $5B_2$ (distal to the point where it sends a process to the RU muscle to avoid stimulation of the excitatory inputs to this muscle) as well as metathoracic nerves 3 and 4 (see Fig. 1) usually evoked either a small depolarisation or hyperpolarisation in the red but not in the white fibres (Fig. 1) (C. Walther, personal communication). The properties of this response were identical to those of inhibitory postsynaptic potentials (i.p.s.ps) found in other locust muscles¹⁰; for example, it was reversibly abolished in Cl-free saline, and it had a reversal potential within a few millivolts of the resting potential (-50 to -60 mV in 10 mM K, locust saline¹⁰) (Fig. 1). Although i.p.s.ps could sometimes be seen in red fibres only when the membrane potential was displaced by ± 5 mV, they were never recorded from white fibres even when the membrane potential was altered by ± 15 mV.

Topically applied GABA (10^{-4} – 10^{-3} M) caused $\sim 50\%$ reduction in the input resistance of the red fibres and also abolished the i.p.s.ps. Picrotoxin (5×10^{-4} to 10^{-3} M) reversibly abolished the i.p.s.ps and blocked the change in input resistance caused by GABA. Responses to GABA ejected iontophoretically from one barrel of a double-barrelled micropipette could be obtained from those regions in the clefts between fibres where a large, fast-rising depolarisation to glutamate was obtained from the second barrel, but not from extrajunctional areas where the glutamate sensitivity was very low. The reversal potential of the GABA response was similar to that of the i.p.s.ps. Bath application of GABA caused a significant change in input resistance of some of the white fibres, suggesting that perhaps these fibres

receive an inhibitory input, although i.p.s.ps were not recorded from them.

In part of this study the RU muscle was denervated by sectioning metathoracic nerve 5 using an aseptic procedure⁹. It is possible that, like the RU excitatory axons, the inhibitory innervation to this muscle runs in metathoracic nerve 5, although it probably supplies other muscle fibres in the leg in addition to those of the RU. The RU on one side of each animal was denervated; the contralateral muscle was left unoperated to serve as a control. Operated animals were maintained in separate containers at 30°C and fed fresh grass daily. Severing nerve 5 interrupts both the excitatory and inhibitory inputs to the RU simultaneously, and it is not feasible to section selectively excitatory and inhibitory axons. To selectively remove the excitatory inputs to the muscle, we have developed a procedure which leaves the inhibitory innervation of the RU muscle intact while deleting the excitatory input to this muscle. The procedure involved surgically excising that part of the metathoracic ganglion which contains the cell bodies of the RU excitatory motoneurons (Fig. 1a), but metathoracic nerve 5 was not sectioned. Animals operated in this way survived well for up to 4 weeks, which was the longest period studied.

Excitatory motor nerve terminal degeneration was complete 2–3 d after section of nerve 5, as judged by the absence of spontaneous miniature excitatory postsynaptic potentials and lack of response of the RU muscle to stimulation of the distal stump of nerve 5 (ref. 12). The sensitivity of the extrajunctional muscle membrane to L-glutamate applied iontophoretically through high-resistance (100–200 M Ω , 0.1 M Na-glutamate, pH 8) micropipettes was examined at various times up to 36 d after section of nerve 5 (and sometimes also nerve 3). The extrajunctional glutamate sensitivity increased significantly compared to control levels 4–6 d after nerve section, and reached a peak after 12–15 d, when the extrajunctional glutamate sensitivity was about 30 times greater than that of control muscles. The increase in extrajunctional sensitivity was accompanied by a decline in membrane potential and an increase in the input resistance of the muscle fibres¹³. These changes were similar in the red and white fibres. Animals with operated ganglia were examined 14 or more days post-operatively. By this time excitatory innervation had degenerated and the extrajunctional membrane of the RU muscle exhibited a high

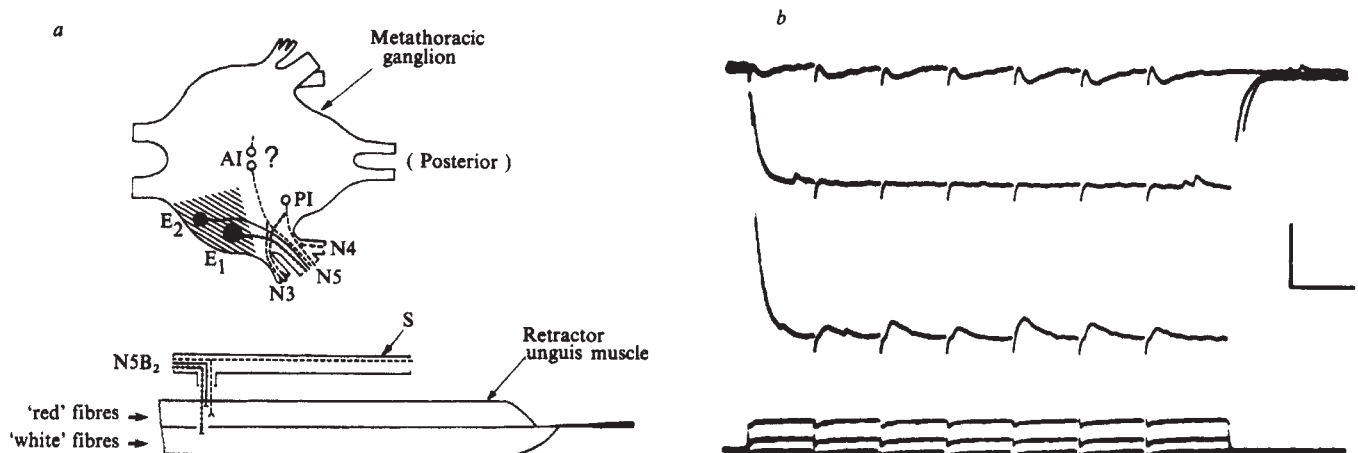


Fig. 1 a, Schematic diagram of the innervation of the femoral portion of the metathoracic retractor unguis (RU) muscle of *Schistocerca gregaria*. Each of the bundles of fibres ('red' and 'white') are innervated by one or two excitatory motoneurons in the metathoracic ganglion (E_1 , E_2). The red fibres are also innervated by one or more inhibitory motoneurons, possibly branches of the anterior (AI) and posterior (PI) inhibitors which innervate the flexor tibiae (FT) muscle. The approximate positions of the cell bodies of E_1 , E_2 , AI and PI are shown (see refs 15 and 16 for details). Axons of both excitatory motoneurons run in metathoracic nerve 5, while axons of inhibitory motoneurons run in nerves 3, 4 and 5. The exact location of the inhibitory axons is not yet clear, and they are represented schematically. In some experiments the shaded part of the metathoracic ganglion was excised to delete selectively the excitatory innervation of the RU muscle. b, Intracellularly recorded inhibitory postsynaptic potentials (i.p.s.ps), evoked in a red fibre of the RU muscle by antidromic stimulation of branch B_2 of nerve 5 (S, in a). This nerve branch also innervates the FT muscle of the metathoracic leg. In this muscle fibre the i.p.s.ps were hyperpolarising at the resting potential, -54 mV. The membrane potential (upper traces) was altered by current injected through a second intracellular microelectrode (lower traces). Note that the i.p.s.ps disappeared when the membrane potential was hyperpolarised to -57 mV. Further hyperpolarisation reversed the polarity of the i.p.s.ps. Membrane potential, 2 mV; current, 20 nA; time, 100 ms. I.p.s.ps could also be recorded in RU fibres by antidromic stimulation of the inhibitory motoneurons through the axons which run in metathoracic nerves 3 and 4.

sensitivity to L-glutamate, similar to that of nerve-sectioned animals.

An example of the increased extrajunctional glutamate sensitivity which occurred on red fibres of denervated RU muscle is shown in Fig. 2. The excitatory innervation of the fibre had degenerated, but i.p.s.ps similar to those recorded from control fibres were recorded from the denervated fibre (Fig. 2). Qualitatively similar results were obtained from red fibres of muscles up to 18 d after nerve section, indicating that up to this time the inhibitory endings on these fibres were still capable of storing and releasing transmitter despite the fact that the axons had been sectioned. Similar results were obtained from animals with an operated ganglion, but in these cases not only could i.p.s.ps be recorded in red fibres on antidromic stimulation of N5B₂, but also during 'spontaneous' and 'reflex' activity of the inhibitory innervation and by stimulation of the inhibitory neurones through metathoracic nerves 3 and 4, indicating that the inhibitory inputs to the RU muscle were intact despite the disappearance of the excitatory innervation.

After periods of more than 21 d after section of nerve 5, i.p.s.ps were absent from all red fibres examined, suggesting that the inhibitory terminals had degenerated, although we have no structural evidence to support this conclusion at present. Electron microscope studies⁴ of the branches of nerve trunk 5B₂ which innervate the RU reveal that the excitatory axons are about 8–10 μm in diameter, whereas other axons in the branch, one or more of which are presumably inhibitory, are about 1 μm in diameter. The disappearance of the excitatory innervation 2–3 weeks before the disappearance of the inhibitory is probably due to more rapid degeneration of the large excitatory axons compared with the small inhibitory axons⁴.

These experiments demonstrate that in red fibres of locust RU muscles, changes in properties resulting from degeneration of excitatory innervation, for example increased extrajunctional L-glutamate sensitivity, decline in membrane potential and increase in membrane resistance, can occur in the presence of a functionally intact inhibitory innervation. It is possible that trophic influences exerted over the red RU fibres by the inhibitory innervation may be relatively weak and nonspecific, as the inhibitory neurone which innervates the RU muscle probably innervates fibres of other muscles in the leg (for example flexor tibiae). In the case of the nerve section experiments any trophic influences exerted by the inhibitory innervation may be rapidly lost after section of nerve 5. The fact that mechanisms responsible for the production and release of inhibitory transmitter apparently remain intact up to about 3 weeks after nerve section gives little information about trophic factors associated with the inhibitory innervation which may be independent of the inhibitory transmitter. For example, trophic materials could be synthesised in the inhibitory neurone(s) and transported to the junctions by a fast component of axonal flow¹⁴. Following nerve section stores of such materials in the distal stump of the inhibitory axons would be rapidly depleted, although the nerve terminals could remain functional in a synaptic transmission sense for some time. However, these objections should not apply to the ganglion operation experiments, where the inhibitory innervation presumably remained intact.

One major difference between the excitatory and inhibitory inputs to the RU muscle lies in the ionic currents gated by the respective neurotransmitters. Whereas i.p.s.ps are generated exclusively by an increase in membrane permeability to Cl, the excitatory transmitter (probably L-glutamate) mediates an inward current of Na and Ca ions¹⁷. In view of the recent findings of Betz and Changeux¹⁸ which suggest that Ca ions, together with cyclic nucleotides, may be important trophic factors in controlling the density of extrajunctional acetylcholine receptors on vertebrate muscle, it is tempting to speculate that the changes in RU muscle properties that result from deletion of the excitatory inputs to the muscle arise indirectly from elimination of the influx of Ca ions into the muscle fibres which accompanies excitatory activity at the nerve-muscle junctions. The Ca ion influx would also be absent even if the inhibitory innervation to

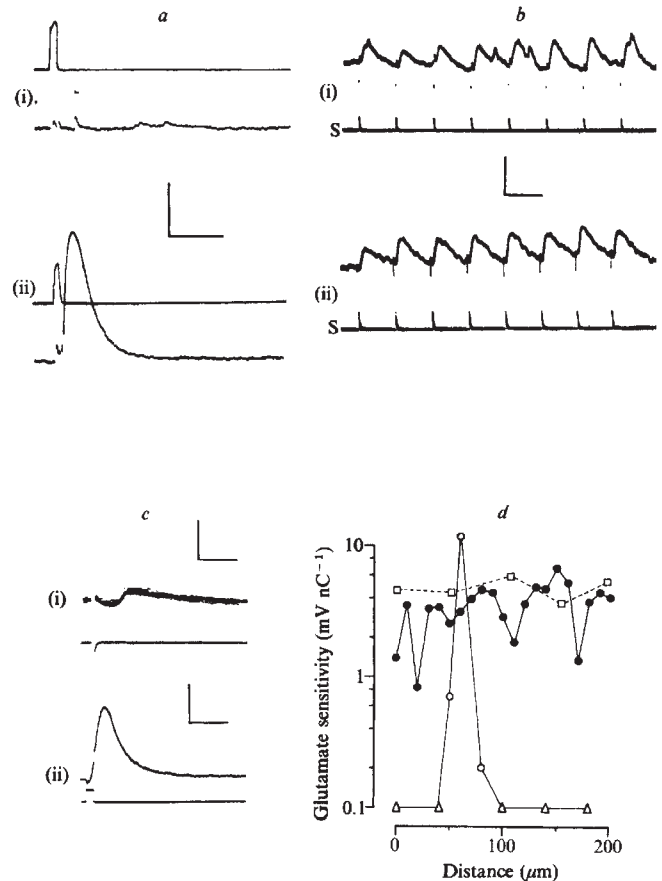


Fig. 2 *a*, Effect of denervation on the L-glutamate sensitivity of RU muscle fibres. (i) A dose of 2 nC (upper trace) applied to the extrajunctional membrane of an innervated control fibre failed to elicit a detectable response (lower trace). (Small responses were elicited with much higher doses of glutamate.) Note miniature excitatory postsynaptic potentials (min. e.p.s.ps). Resting potential of the fibre was -53 mV. (ii) A similar dose of glutamate applied to the extrajunctional membrane of a 'red' fibre of the muscle 16 d after nerve section evoked a large depolarisation. Note absence of min. e.p.s.ps. Resting potential of this fibre was -44 mV. Vertical and horizontal calibration bars, 2 mV, 100 ms, respectively. *b*, (i) I.p.s.ps recorded at the resting potential from the same control fibre as in (a). (ii) I.p.s.ps very similar to those recorded from the control fibre were recorded from the denervated fibre on stimulation of nerve 5B₂ (see Fig. 1). Calibration bars, 0.5 mV, 100 ms, respectively. *c*, (i) Superimposed i.p.s.ps recorded from an RU muscle fibre of an animal 18 d after surgical deletion of the RU excitatory motoneurons. Resting potential of the fibre was -46 mV. These i.p.s.ps were recorded on stimulation of nerve 5B₂, but similar responses were obtained on stimulation of metathoracic nerves 3 and 4. Calibration bars, 1 mV, 20 ms, respectively. (ii) Response to a 1-nC dose of glutamate (bottom trace) applied to the extrajunctional membrane of the same muscle fibre. Note that the excitatory innervation of the muscle had degenerated, that is no min. e.p.s.ps were observed on this or any other record from this fibre and e.p.s.ps. could not be evoked from the RU muscle fibres during stimulation of metathoracic nerve 5. Calibration bars, 2 mV, 200 ms, respectively. *d*, Comparison of the distribution of glutamate sensitivity of an innervated (Δ , $\circ$) and a denervated ($\bullet$) RU muscle fibre. The glutamate electrode was moved in 10–20- μm steps along the surface of the muscle fibre and the sensitivity to glutamate at each point measured in terms of the peak depolarisation produced by a 2-nC dose of glutamate. The sensitivity of the innervated fibre was less than 0.1 mV nC⁻¹ (Δ) at most points on the fibre; the localised area of high sensitivity ($\circ$) corresponds to a neuromuscular junction. By contrast, extrajunctional membrane of the denervated fibre (16 d after section of nerve 5) had a large and nearly uniformly distributed glutamate sensitivity although not as high as that of former junctional sites (not shown). ($\square$), Illustrates the extrajunctional glutamate sensitivity of an RU muscle fibre 16 d after deletion of the excitatory motoneurons.

the muscle remained intact, as Ca ions do not contribute to the inhibitory currents.

Further studies are in progress to evaluate the effect of selective deletion of inhibitory innervation on the properties of RU muscle.

R. B. CLARK
K. A. F. GRATON
P. N. R. USHERWOOD

Department of Zoology,
University of Nottingham,
Nottingham, UK

Received 19 April; accepted 6 July 1979.

1. Usherwood, P. N. R. & Cull-Candy, S. J. in *Insect Muscle* (ed. Usherwood, P. N. R.) 207-280 (Academic, London, 1975).
2. Harris, A. J. A. *Rev. Physiol.* **36**, 251-305 (1974).
3. Gutmann, E. A. *Rev. Physiol.* **38**, 177-216 (1976).
4. Rees, D. & Usherwood, P. N. R. *Comp. Biochem. Physiol.* **43A**, 83-101 (1972).
5. Usherwood, P. N. R. *Nature* **223**, 411-413 (1969).
6. Cull-Candy, S. G. *Nature* **258**, 530-531 (1975).
7. Cull-Candy, S. G. *J. Physiol., Lond.* **276**, 165-181 (1978).
8. Clark, R. B., Graton, K. A. F. & Usherwood, P. N. R. *J. Physiol., Lond.* **290**, 551-568 (1979).
9. Usherwood, P. N. R. *J. Insect Physiol.* **9**, 247-255 (1963).
10. Usherwood, P. N. R. & Grundfest, H. *J. Neurophysiol.* **28**, 497-518 (1965).
11. Cochrane, D. G., Elder, H. Y. & Usherwood, P. N. R. *J. Cell Sci.* **10**, 419-441 (1972).
12. Usherwood, P. N. R. *J. exp. Biol.* **59**, 1-16 (1973).
13. Graton, K. A. F., Clark, R. B. & Usherwood, P. N. R. *Neuropharmacology* **18**, 201-208 (1979).
14. Ochs, S. *Ann. N.Y. Acad. Sci.* **228**, 202-223 (1974).
15. Burrows, M. & Horridge, G. A. *Phil. Trans. R. Soc. B* **269**, 49-84 (1974).
16. O'Shea, M. & Fraser Rowell, C. H. in *Identified Neurons and Behaviour in Arthropods* (ed. Hoyle, G.) 307-328 (Plenum, New York, 1977).
17. Anwyl, R. & Usherwood, P. N. R. *Nature* **252**, 591-593 (1974).
18. Betz, H. & Changeux, J. P. *Nature* **278**, 749-752 (1979).

Differential labelling of retinal neurones by ^3H -2-deoxyglucose

THE delineation of structural and functional interconnections in the vertebrate retina has long been a central problem in vision research. The Golgi technique^{1,2}, combined with ultrastructural studies^{3,4} and post-recording injection of intracellular stains^{5,6}, has revealed the morphological and physiological properties of many of the individual cell types of the retina. Recently, light and electron microscope autoradiography with putative neurotransmitters has tentatively identified certain retinal neurones with specific transmitters^{7,8}. However, very few techniques exist which can reveal the interconnections of a 'pathway' of retinal neurones in a functionally dependent manner. Radioactive 2-deoxyglucose (2-DG) is commonly used as a tracer in the central nervous system, where biochemical and autoradiographic studies have shown it to be a good indicator of local functional activity⁹⁻¹¹. Transport and phosphorylation of 2-DG are identical to those of glucose, but no further metabolism beyond 2-DG-6-PO₄ occurs¹². Most published studies using 2-DG as a tracer have used the ^{14}C -labelled form and frozen thick-sectioned tissue, a combination in which the resolution of morphological structure is unacceptable for studying neuronal connections in the retina. In the studies reported here, we have used tritium-labelled 2-DG in combination with plastic embedding to produce stimulus-dependent labelling at the cellular level in the isolated goldfish retina; our results suggest that these modifications of the '2-deoxyglucose technique' are useful for studying the physiology and functional connections of retinal neurones.

We chose the retina of the goldfish (*Carassius auratus*) for these studies because of its well characterised morphology and physiology^{1,3-6}. After overnight dark-adaptation, goldfish eyecups were removed using an IR image converter. Small pieces of isolated retinas were removed from the eyecups under the same illumination as was used during incubation. For dark incubation, only IR illumination was used. Retinas were incubated in 0.2-0.5 ml of teleost Ringer's solution containing 100 mM NaCl, 3.5 mM KCl, 1.5 mM MgCl₂, 0.2 mM NaH₂PO₄, 30 mM

NaHCO₃, 1.4 mM CaCl₂ and 2 mM glucose, pH 7.3, which had been presaturated with 95% O₂/5% CO₂. 2-Deoxy-D-[1- ^3H]glucose (Amersham, TRK 383) was present at 1 mCi ml⁻¹.

Incubations were carried out in small vials sealed with cover slips and centred under the photostimulator beam. During incubation, retinas were maintained in the dark or stimulated with flashing broad-band white light or narrow-band red light of equal quanta. Stimulation was provided by a 6-V tungsten source with a maximum output of 300 $\mu\text{W cm}^{-2}$. The source gave a 22-mm diameter beam of collimated light. Neutral density filters were used to attenuate the light to 2×10^{12} quanta s⁻¹ cm⁻², calculated at 550 nm (average 400-700 nm) for broad-band white light and isolated at 700 nm with an interference filter having a 13-nm half-peak bandwidth. This stimulus was delivered as a 1-Hz, 50% duty cycle square wave through the use of a rotating disk. After 20 min, an equal volume of double-strength fixative was added to the incubation mixture and stimulation continued for an additional 15 min. The pieces were

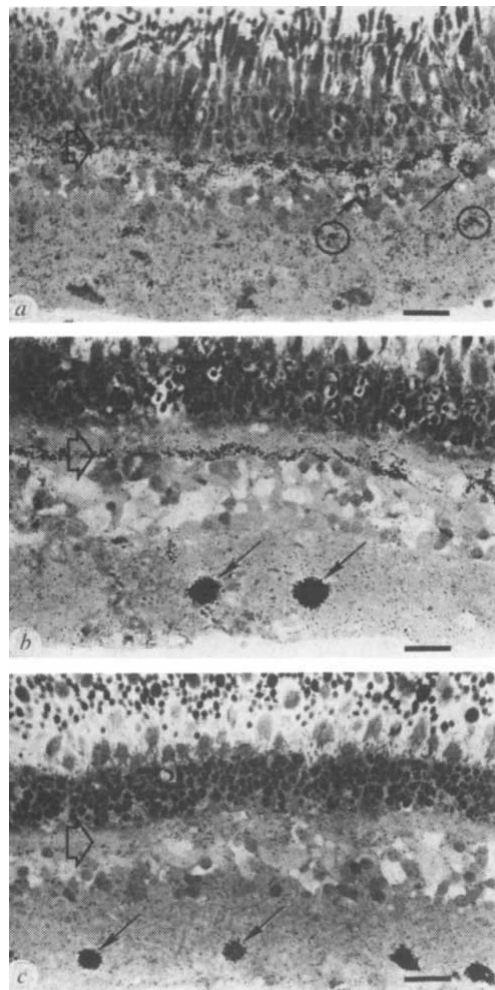


Fig. 1 Light microscope autoradiographs of isolated goldfish retinas incubated for 20 min in teleost Ringer's solution containing 1 mCi ml⁻¹ ^3H -2-DG. *a*, Incubated in darkness after IR dissection. Label appears predominantly in photoreceptors, H1 horizontal cells (open arrow) and presumptive type *a* (centre-hyperpolarising) bipolar cells (closed arrows). The heavily labelled areas just beneath these cells, in sublamina *a* of the inner plexiform layer, are probably their axon terminals (open circles). *b*, Incubated under flashing narrow-band red light (700 nm) at 2×10^{12} quanta s⁻¹ cm⁻². Label appears in H2 and/or H3 horizontal cells (open arrow) and type *b* (centre-depolarising) bipolar cell axon terminals (closed arrows). *c*, Incubated under flashing broad-band white light at 2×10^{12} quanta s⁻¹ cm⁻². Very little label is seen in photoreceptors and horizontal cell layers (open arrow); only type *b* bipolar cell axon terminals are heavily labelled (closed arrows). Scale bars, 20 μm .

transferred to fresh 3% glutaraldehyde for 1–2 h, then rinsed twice for 10 min in 0.1 M phosphate buffer, pH 7.3, post-fixed for 1 h in buffered 1% OsO₄, rinsed twice for 10 min in buffer, dehydrated in graded ethanol and acetone, then embedded in Epon-araldite and polymerised for 48 h at 65 °C. Longitudinal and tangential sections of 1 µm were transferred to glass slides, coated with Kodak NTB2 emulsion and exposed at 4 °C for 2–3 weeks. After exposure, the slides were developed in Dektol for 2 min at 17 °C, washed, then stained with 1% toluidine blue. Similar retinas which were not embedded in plastic, but were extracted and analysed by paper chromatography¹³, revealed that more than 85% of the tracer was present as 2-DG-6-PO₄, the remainder being 2-DG.

One of the major aims of this study was to determine whether ³H-2-DG-labelled retinas embedded in plastic would provide resolution at the single cell level¹⁴. Figure 1 shows an autoradiograph from an experiment in which dark-adapted goldfish retinas were incubated with ³H-2-DG in darkness (a), under flashing red light (700 nm) (b), and under flashing white light (c). This figure demonstrates that our modification of the 2-DG technique can be used to obtain resolution at the cellular level. Additionally, although we have not yet completed a detailed analysis of all the labelled cell types in these retinas, our preliminary analysis reveals several stimulation-dependent differences in the labelling patterns, possibly reflecting the functional dependence of ³H-2-DG labelling. Dark-adapted retinas incubated in the dark with ³H-2-DG (Fig. 1a) show considerable labelling in most rod and cone photoreceptors, heavy labelling in H1 horizontal cells (luminosity- or L-type) but not other horizontal cells, and heavy labelling in some presumptive type *a* (centre-hyperpolarising) but not type *b* (centre-depolarising) bipolar cells (terminology used is that of refs 4 and 6). Retinas incubated under flashing red light (Fig. 1b) show less labelling in photoreceptors, little labelling in H1 horizontal cells, but heavy labelling in H2 and/or H3 horizontal cells (chromaticity- or C-type) and in some type *b* bipolar cells. Retinas incubated under flashing white light (Fig. 1c) show almost no labelling in photoreceptors, only diffuse labelling in all horizontal cells, and heavy labelling in some type *b* bipolar cells. In all cases, very few amacrine or ganglion cells are labelled, possibly because the mode of stimulation was not optimal for labelling these cells.

In many cases, single 1-µm longitudinal sections through the retina were not very useful for identifying individual cell types. Accordingly, tangential and serial longitudinal sections were cut, prepared for autoradiography and exposed for longer periods of time. Figure 2 shows that these techniques permit the identification of specific cells in the retina. Tangential sections through the inner nuclear layer of the goldfish retina reveal the morphology of the horizontal cells which have taken up ³H-2-DG. Figure 2a shows an H1 horizontal cell identified by this method, and for comparison, Fig. 2b shows an isolated H1 horizontal cell obtained by enzymatic dissociation¹⁵. Similar identifications can be made for other retinal neurones. To identify bipolar cells, autoradiograms of serial 1-µm longitudinal sections were traced in sequence and the tracings superimposed to demonstrate the cell's morphology (Fig. 2c). When these reconstructions are compared with isolated cells (Fig. 2d) or with cells impregnated by the Golgi technique (Fig. 2e) they can be readily identified as bipolar cells. In this particular case, the functional, positional and morphological characteristics reveal these cells to be type *b* bipolar cells^{3–6}. The identification of the other labelled retinal neurones by these methods is now in progress. Thus, ³H-2-DG incorporation under specific stimulus conditions can be used to demonstrate the functional morphology of certain retinal cell types.

Increased 2-DG labelling has been shown to reflect increased functional activity of neural tissue¹⁶. For retinal neurones, our autoradiographic results are consistent with the idea that accumulation is influenced by the membrane potentials of the cells, with depolarisation leading to enhanced labelling and hyperpolarisation to decreased labelling. The strongest evidence for

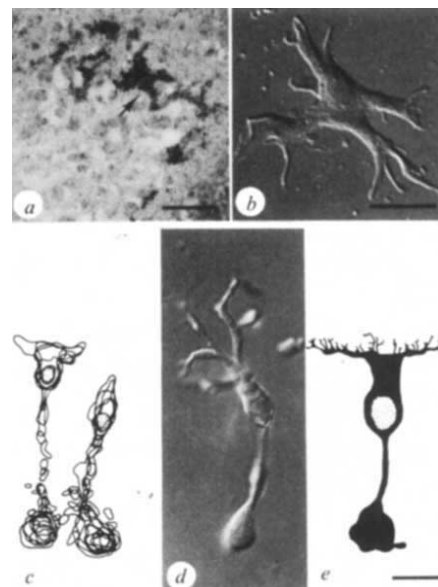


Fig. 2 Identification of ³H-2-DG-labelled cells in the goldfish retina. *a*, Tangential section through the inner nuclear layer after dark incubation with ³H-2-DG and subsequent autoradiographic preparation. An H1 horizontal cell is filled with label (arrow). Scale bar, 30 µm. *b*, For comparison, an isolated H1 horizontal cell prepared according to the dissociation technique of Lam¹⁵. Scale bar, 20 µm. *c*, Camera lucida tracings of serial longitudinal sections of a retina labelled under red light were superimposed to produce reconstructions of type *b* bipolar cells with axon terminals in sublamina *b* of the inner plexiform layer. Stippled areas identify the nuclei. *d*, An isolated type *b* bipolar cell prepared by cell dissociation. *e*, A type *b* bipolar cell impregnated by the Golgi technique (redrawn from Stell *et al.*⁴). Scale bar for *c*, *d* and *e*, 10 µm.

this suggestion is the intense labelling of type *b* bipolar cells under red and white stimulation, conditions which are known to depolarise these cells. Further evidence lies in the increased labelling of H1 horizontal cells and presumptive type *a* bipolar cells in the dark, conditions in which they are depolarised; the H1 cells accumulate little label under red stimulation, a condition in which they are hyperpolarised. Thus, it seems that in all three incubation conditions, neurones in a depolarised state accumulate more label than those in a hyperpolarised state. This method may therefore be useful for the elucidation of neuronal connections and functional architecture in the retina, especially in retinas not easily amenable to intracellular recordings, as well as in other areas of the central nervous system.

This investigation was supported by grants from the Retina Research Foundation, the Retinitis Pigmentosa Foundation and grants EY 01406 and EY 02423 from NIH. S.F.B. and D.M.K.L. are recipients of Research Career Development Awards from the National Eye Institute, and W.C.G. is an NRSA postdoctoral fellow. We thank R. Hoffman and L. Marroquin for technical assistance, C. Roeder-Gordon for preparation of the manuscript, R. Benolken for help with the photostimulator, and R. Marc and H. Sperling for helpful discussions and use of the IR converter.

SCOTT F. BASINGER
WILLIAM C. GORDON
DOMINIC M. K. LAM

Cullen Eye Institute,
Baylor College of Medicine,
Houston, Texas 77030

Received 20 March; accepted 2 July 1979.

1. Cajal, S. R. *The Structure of the Retina* (Thomas, Springfield, 1972).
2. Boycott, B. B. & Dowling, J. E. *Phil. Trans. R. Soc. B* **255**, 109–184 (1969).
3. Stell, W. K. & Lightfoot, D. O. *J. comp. Neurol.* **159**, 473–501 (1975).

4. Stell, W. K., Ishida, A. T. & Lightfoot, D. O. *Science* **198**, 1269–1271 (1977).
5. Kaneko, A. *J. Physiol. Lond.* **207**, 623–633 (1970).
6. Famiglietti, E. V., Kaneko, A. & Tachibana, M. *Science* **198**, 1267–1269 (1977).
7. Bonting, S. L. *Transmitters in the Visual Process* (Pergamon, New York, 1976).
8. Marc, R., Stell, W. K., Bok, D. & Lam, D. M. K. *J. comp. Neurol.* **182**, 221–245 (1978).
9. Reivich, M., Sano, N. & Sokoloff, L. in *Brain and Blood Flow* (ed. R. W. Ross-Russell) (Pittman, London, 1971).
10. Hubel, D. H., Wiesel, T. N. & Stryker, M. P. *Nature* **269**, 328–330 (1977).
11. Des Rosiers, M. H. *et al. Science* **200**, 447–449 (1978).
12. Kennedy, C. *et al. Science* **187**, 850–853 (1975).
13. Schwartz, R. T. & Schmidt, F. G. *Eur. J. Biochem.* **62**, 181–187 (1976).
14. Des Rosiers, M. H. & Descarries, L. *C. r. hebdom. Séanc. Acad. Sci., Paris D287*, 153–156 (1978).
15. Lam, D. M. K. *Cold Spring Harb. Symp. quant. Biol.* **40**, 571–579 (1975).
16. Sokoloff, L. *et al. J. Neurochem.* **28**, 897–916 (1977).

Glucocorticoids inhibit endorphin synthesis by pituitary cells

ENDORPHINS in the pituitary gland^{1,2} are specifically localised in the intermediate and anterior lobes^{3,4}, whereas the posterior lobe contains much less opioid activity⁴. Recent results have indicated that β -endorphin and adrenocorticotrophic hormone (ACTH) are synthesised by pituitary cells from a common precursor polypeptide^{5–8} and that they both appear in the circulation after stressful stimulation⁹. It is well known that an increased potassium concentration stimulates the release of ACTH from pituitary cells in a calcium-dependent fashion¹⁰. Similar potassium-evoked release of endorphins was indicated in AtT/20 cells which originated in the anterior lobe of mouse pituitary tumours¹¹. In addition, a basal release of endorphins was indicated, suggesting a possible dual regulation of endorphin secretion from the pituitary¹¹. Coordinated release of ACTH, β -endorphin and β -lipotropin was reported¹² while the present study was in progress. This study was undertaken to determine whether glucocorticoid hormones, which are known to regulate ACTH, also regulate pituitary endorphins. Although it has recently been shown⁹ that dexamethasone inhibits *in vivo* elevation of the content of pituitary endorphins after adrenalectomy, the mechanism of the action of this synthetic glucocorticoid remains unclear. Here, we describe data suggesting a direct inhibitory role of glucocorticoids in the synthesis of endorphins by pituitary cells.

The results shown in Fig. 1a indicate that the content of endorphins in AtT/20 pituitary cells was decreased when the cells were cultured with the synthetic glucocorticoids, dexamethasone and prednisolone. This effect is dose dependent and shows half-maximal response at 1.1 ± 0.5 nM and 1.6 ± 0.6 nM for dexamethasone and prednisolone, respectively. The inhibitory role of the glucocorticoids was observed at these physiological concentrations only when the serum used for cell culture was extensively dialysed (Fig. 1). If non-dialysed serum was used, dexamethasone had significantly lower activity, decreasing the cellular content of endorphins by only $35 \pm 7\%$ and $60 \pm 12\%$ with 100 and 1,000 nM, respectively. These findings agree with previous studies which showed that the effect of steroids could be observed at physiological concentrations only after dialysis or charcoal absorbance, apparently through their removal of free steroids or some other form of inhibition.

As AtT/20 cells release endorphins spontaneously^{11,12}, we needed to know whether the decreased cellular endorphin content observed after treatment with prednisolone or dexamethasone reflected a shut-off of the synthesis of these peptides or whether the hormones caused the fall in the cellular levels by potentiating the release system. Therefore, the level of endorphins in the culture medium was also determined. Figure 1b shows that concomitantly with the drop in the cellular endorphin content seen with increasing concentrations of prednisolone or dexamethasone, there was a dose-dependent fall in the amount of endorphins in the culture medium. The effect of dexamethasone (10 nM) on the release of endorphins at different days after addition of the hormone showed a simple

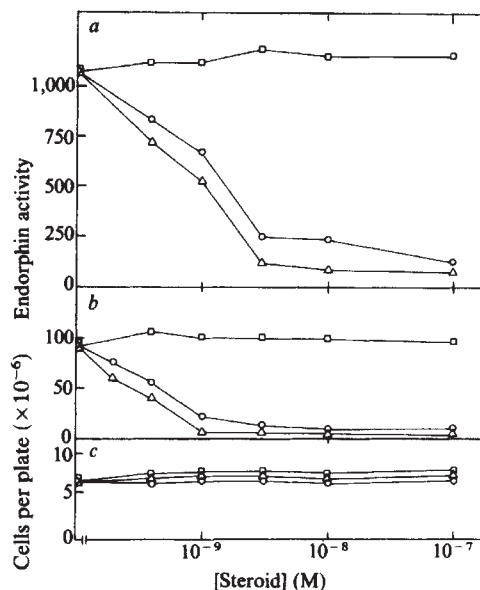


Fig. 1 Effect of steroid hormones on cellular and released endorphins from pituitary cells. AtT/20 mouse pituitary cells were cultured as described previously¹¹ except that the fetal calf serum (H-21, GIBCO) was extensively dialysed for 4 consecutive days against 50 volumes (each day) of 0.15 M NaCl. Cells were seeded for the experiments at 5×10^5 cells in 5 ml of culture medium containing 10% dialysed serum in a 60-mm Petri dish (Nunc). The cultures were treated on the day of inoculation and 48 h later with 5 μ l of ethyl alcohol containing oestradiol benzoate ($\square$), prednisolone ($\circ$) or dexamethasone ($\triangle$) to give the indicated final hormone concentrations. Control dishes obtained 5 μ l of ethyl alcohol. 48 h after the last hormone addition the cells were collected with the culture medium and the plates washed with 3 ml of phosphate-buffered saline, pH 7.4, which was added to the media containing the cells. This mixture was divided into two portions and centrifuged for 3 min in a low-speed centrifuge (500g). The medium of one portion was collected and stored at -20°C and the cells were suspended in 0.1% EDTA in phosphate-buffered saline (PBS) and counted (C). The medium of the other portion was discarded and the cells were extracted in 3 ml of 0.1 M HCl. The homogenate was neutralised with 1/5 volume of 0.5 M Tris base, centrifuged for 10 min at 40,000g and the supernatant stored at -20°C . Endorphin activity of the cell extracts (a) or the media (b) was quantitated by the opiate radioreceptor assay. We used this method rather than radioimmunoassay for quantitation of endorphins to ensure that all the endorphins were detected and also to avoid any possibility of misinterpretation due to cross-reaction between endorphin antibodies and other pituitary peptides. In these conditions, β -lipotropin was inactive. Rat brain tissue was used as the source of opiate receptors and details of the assay are described elsewhere¹¹. Aliquots of 50–100 μ l of the AtT/20 cell extracts or culture media were tested in duplicate and their activity was calculated from standard curves of Met-enkephalin (1–500 nM). Endorphin activity is expressed as pmol Met-enkephalin equivalents per 10^6 cells. Data presented are the average of three experiments carried out with duplicate plates for each hormone treatment.

time-dependent drop with no indication of any increases (data not shown). It was possible that the decreased endorphin levels in the presence of the glucocorticoids were a result of inhibition or other modification in cell proliferation. However, this was ruled out because (1) cells of each plate were counted and the cellular as well as the released endorphins were calculated per cell, and (2) there were no significant changes in cell multiplication after treatment with 0.2–100 nM prednisolone or dexamethasone (Fig. 1c).

The specificity of the effect of the steroids on AtT/20 cell endorphin content was also studied. Figure 1 indicates that 0.4–100 nM of Oestradiol benzoate had no significant effect on either the cellular content or the amounts of endorphins released. Two other sex hormones, progesterone and testosterone, and two other steroids, 17- α -epitestosterone and Δ^4 -androstene-11 β -3,17-dione, had no effect at 10 and 100 nM (Table 1). The specificity of the effect of the glucocorticoids was further indicated by the fact that incorporation of radiolabelled ^3H -dexamethasone into the nuclei of AtT cells was selectively inhibited by dexamethasone or prednisolone at nanomolar

concentrations. In the same conditions, oestradiol benzoate, testosterone, 17- α -epitestosterone and Δ^4 -androstene-11 β -3,17-dione at 10 or 100 nM did not inhibit ^3H -dexamethasone incorporation by these cells (Table 1).

The inhibitory role of adrenal glucocorticoids on ACTH secretion seems to be mediated both by inhibition of the release of hypothalamic corticotropin-releasing factors and by a direct effect on the pituitary cells (ref. 13 and refs therein). The present study suggests that adrenal glucocorticoids at physiological concentrations inhibit endorphin synthesis by anterior pituitary AtT/20 cells in cultures that lack added hypothalamic factors. A previous study with AtT/20 cells¹² suggested that dexamethasone inhibits the stimulatory effect of hypothalamic factors which cause an increased release of endorphins. However, as the cells were cultured for 48 h with a high dexamethasone concentration (1,000 nM), one should reconsider the possibility that dexamethasone acted by decreasing the cellular endorphin content. This is supported by studies which indicated a concomitant decrease of cellular content and the basal release of ACTH after treatment of AtT cells with dexamethasone¹⁴.

Recently, we have observed fluctuations in the content of pituitary endorphins during the normal oestrous cycle of female rats¹⁵. Interestingly, the lowest level of pituitary endorphins occurred at the day of proestrus, which was also found to be the day in which there was release of adrenal corticosterones¹⁶. It is conceivable, therefore, that physiological and experimental conditions that increase glucocorticoid levels in the blood can inhibit the synthesis of endorphins by pituitary cells and thereby decrease the amount of endorphins available for release to the peripheral circulation or other target tissues. Acute injection of morphine to experimental animals stimulates ACTH and adrenal hypersecretion. The fact that removal of the source of glucocorticoids in the body (by adrenalectomy) renders rats more susceptible to morphine is well known¹⁷. Thus, the possibility that glucocorticoid-induced suppression of endorphin synthesis in the pituitary gland may fulfil a role in chronically treated subjects deserves further study.

Table 1 Specificity of the effect of various steroids on cellular content of endorphins and on incorporation of ^3H -dexamethasone to AtT/20 pituitary cells

Steroid	Concentration (nM)	Cellular endorphins	^3H -Dexamethasone incorporation (fmol per 10^7 cells)
None	—	1,043 $\pm$ 165	28.4 $\pm$ 4.6
Dexamethasone	0.4	725 $\pm$ 110*	NT
	1.0	520 $\pm$ 85*	26.5 $\pm$ 3.6
	10.0	92 $\pm$ 60*	23.8 $\pm$ 1.1†
	100.0	80 $\pm$ 35*	16.4 $\pm$ 0.1*
Prednisolone	0.4	835 $\pm$ 115†	NT
	1.0	670 $\pm$ 100*	27.2 $\pm$ 5.1
	10.0	198 $\pm$ 60*	24.7 $\pm$ 3.2†
	100.0	105 $\pm$ 55*	17.5 $\pm$ 0.2*
Oestradiol benzoate	10.0	1,147 $\pm$ 190	29.7 $\pm$ 3.6
	100.0	1,260 $\pm$ 220	23.4 $\pm$ 4.4
Progesterone	10.0	1,061 $\pm$ 160	28.4 $\pm$ 5.5
	100.0	985 $\pm$ 260	27.6 $\pm$ 3.2
Testosterone	10.0	1,140 $\pm$ 117	25.7 $\pm$ 4.0
	100.0	1,031 $\pm$ 125	33.6 $\pm$ 5.5
17- α -Epitestosterone	100.0	1,275 $\pm$ 178	27.7 $\pm$ 3.1
Δ^4 -Androstene-11 β -3,17-dione	100.0	1,074 $\pm$ 125	27.1 $\pm$ 3.7

Cellular endorphin content was determined as described in Fig. 1. Data are pmol Met-enkephalin equivalents per 10^6 cells. Studies of the incorporation of ^3H -dexamethasone (23 Ci mmol⁻¹, Radiochemical Centre) to nuclei of AtT/20 cells were carried out as described elsewhere¹⁸, with minor modifications. Cells were dispersed with 0.1% EDTA in PBS, counted, washed with culture medium without serum and dispersed at 2×10^7 cells per ml of the same medium. The cells were then incubated for 40 min at 37°C in a humidified incubator with 4×10^{-8} M ^3H -dexamethasone with or without the indicated concentration of unlabelled steroids. Nuclei were prepared and radioactivity was determined¹⁸. 10 μM dexamethasone was taken as the baseline of maximal displacement. All data are the average $\pm$ s.e. of duplicate samples from 2–4 experiments carried out independently. NT, Not tested.

* $P < 0.001$; † $P < 0.01$ compared with control samples.

I thank Mr A. Goldenberg for technical assistance and Professor L. Sachs for support and interest. R.S. is an incumbent of the CS Koshland Career Development Chair.

RABI SIMANTOV

Department of Genetics,
Weizmann Institute of Science,
Rehovot, Israel

Received 26 April; accepted 13 July 1979.

- Li, C. H. & Chung, D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1145–1148 (1976).
- Goldstein, A. *Science* **193**, 1081–1086 (1976).
- Bloom, F. *et al. Life Sci.* **20**, 43–48 (1977).
- Simantov, R. & Snyder, S. H. *Brain Res.* **124**, 178–184 (1977).
- Mains, R. E., Eipper, B. A. & Ling, N. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3014–3018 (1977).
- Roberts, J. L. & Herbert, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4826–4830 (1977).
- Rubinstein, M., Stein, S. & Udenfriend, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 669–671 (1978).
- Seidah, N. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3153–3157 (1978).
- Guillemin, R. *et al. Science* **197**, 1367–1369 (1977).
- Kracier, J., Milligan, J. V., Gosbee, J. L., Conrad, R. G. & Branson, C. M. *Endocrinology* **85**, 1144–1153 (1969).
- Simantov, R. *Life Sci.* **23**, 2503–2508 (1978).
- Allen, R. G., Herbert, E., Hinman, M., Shibuta, H. & Pert, C. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4972–4976 (1978).
- Russell, S. M., Dhariwal, A. P. S., McCann, S. M. & Yates, F. E. *Endocrinology* **85**, 512–521 (1969).
- Watanabe, H., Nicholson, W. E. & Orth, D. R. *Endocrinology* **93**, 411–416 (1973).
- Simantov, R., Hershkowitz, M. & Szechtman, H. *Brain Res.* (submitted).
- Raps, D., Barthe, P. L. & Desautel, P. A. *Experientia* **27**, 339–340 (1971).
- Mackay, E. M. & Mackay, L. L. *J. Pharmac. exp. Ther.* **35**, 67–74 (1979).
- Sibley, C. H. & Tomkins, G. *Cell* **2**, 221–227 (1974).

Continuous culture of T cells cytotoxic for autologous human leukaemia cells

CYTOTOXIC T lymphocytes (CTLs), capable of lysing autologous and HLA-identical human leukaemia cells, can be generated *in vitro* by culturing lymphocytes from patients in remission from leukaemia with X-ray irradiated autologous leukaemia cells and irradiated allogeneic normal cells in 'three-cell' mixed leukocyte culture (MLC) protocols^{1–3}. Additionally, T cells cultured with X-ray irradiated normal lymphocytes pooled from 20 unrelated individuals can differentiate into CTLs capable of lysing autologous Epstein-Barr (EB) virus-transformed lymphoblastoid cell lines and autologous human leukaemia cells, but not autologous normal cells^{4,5}. Because *in vitro* generated CTLs directed against syngeneic tumour cells can inhibit tumour development in some animal tumour systems^{6–9}, it is likely that if large numbers of human CTLs cytotoxic for autologous malignant cells could be generated *in vitro*, these CTLs could be effective in eliminating residual malignant cells in patients who have received chemotherapy. Recent findings that normal human T cells^{10,11} and allosensitized mouse¹² and human CTLs^{13,14} grow continuously in medium containing T-cell growth factor (TCGF) derived from supernatants of mitogen-stimulated lymphocytes, have raised the possibility that large numbers of *in vitro* generated CTLs cytotoxic for autologous human malignant cells could be propagated by growing CTLs in TCGF. We report here that *in vitro* generated CTLs cytotoxic for autologous lymphoblastoid cell lines (LCL) and human leukaemia cells can be increased in number by several millionfold by growth in TCGF and that the continuously growing CTLs are highly cytotoxic for autologous abnormal cells.

Experiments were undertaken to determine whether large numbers of T cells cytotoxic for autologous EB virus-transformed LCL cells could be obtained by culturing T cells that had been sensitized to pooled allogeneic normal cells (pool₀) in TCGF. T lymphocytes stimulated with the 'pool' for 7 days lysed autologous LCL cells but not autologous normal cells; however, both the LCL cells and normal lymphocytes were lysed by pool-sensitized T cells from unrelated individual A (Table 1). After 27 days growth in TCGF, the CTLs had increased in number by nearly 10^7 -fold; the resulting cells were at least threefold more cytotoxic for autologous LCL cells than were

Table 1 Lysis of autologous LCL cells by continuous cultures of T cells sensitised to pooled allogeneic normal cells

Effector cells	Days after sensitization	Days grown in TCGF	Effector: target ratio	% Specific ⁵¹ Cr release $\pm$ s.d.			
				J's LCL cells	J's normal lymphocytes	H's LCL cells	H's normal lymphocytes
J pool _x	7	0	30:1	20.0 $\pm$ 5.2	0.9 $\pm$ 5.4		
			10:1	11.8 $\pm$ 1.6	0.4 $\pm$ 4.8		
	34	27	30:1	53.4 $\pm$ 5.5	-4.3 $\pm$ 3.9		
			10:1	34.0 $\pm$ 1.5	-6.4 $\pm$ 3.8		
J J _x	34	27	30:1	13.9 $\pm$ 2.4	-4.0 $\pm$ 4.2		
			10:1	4.1 $\pm$ 1.6	-5.2 $\pm$ 3.8		
H pool _x	7	0	30:1			19.0 $\pm$ 9.3	-4.3 $\pm$ 4.1
			10:1			15.4 $\pm$ 1.9	-3.3 $\pm$ 4.1
	34	27	30:1			29.4 $\pm$ 7.0	-1.8 $\pm$ 2.0
			10:1			22.7 $\pm$ 3.4	-2.5 $\pm$ 1.8
H H _x	34	27	30:1			9.5 $\pm$ 5.0	-1.8 $\pm$ 2.1
			10:1			3.5 $\pm$ 2.9	-4.7 $\pm$ 2.3
A pool _x	7	0	30:1	71.0 $\pm$ 2.1	58.8 $\pm$ 6.1	60.7 $\pm$ 5.4	41.1 $\pm$ 2.1

T cells from the peripheral blood of normal individuals J, H and A were isolated by rosetting Ficoll-Hypaque purified mononuclear cells with sheep red blood cells (SRBC) by a method reported elsewhere¹⁶ which was modified as follows. Briefly, 3 ml of 0.5% SRBC in phosphate-buffered saline (PBS) were mixed with an equal volume of mononuclear cells (5×10^6 cells ml⁻¹) and 0.6 ml heat-inactivated pooled human AB serum which had previously been adsorbed with SRBC as described previously¹⁶. The cell mixture was incubated at 37°C for 5 min, centrifuged at 50g for 5 min and incubated at 4°C for 60 min. The cells were gently resuspended, layered on 10 ml of ice-cold Ficoll-Hypaque and centrifuged for 20 min at 400g. The pellet of rosetted cells was resuspended and SRBC were lysed by hypotonic shock. The T cells were washed with PBS and 5×10^6 responding T cells were cultured with equal numbers of X-ray irradiated (2,500 R) allogeneic lymphocytes pooled from 20 normal individuals (designated pool_x) or with equal numbers of X-ray irradiated autologous normal lymphocytes (designated J_x and H_x) in 50 ml Falcon tissue culture flasks containing 5 ml of RPMI-1640 medium supplemented with 25 mM HEPES buffer and 20% heat-inactivated normal human serum, as previously described⁴. On day 7 after the onset of MLC, some of the effector cells were seeded in 50 ml Falcon tissue culture flasks at 1×10^5 cells ml⁻¹ in 10 ml RPMI-1640 medium supplemented with 25 mM HEPES buffer, 20% heat-inactivated normal human serum and 10% TCGF (Associated Biomedical Systems). Before use TCGF was concentrated fourfold against polyethylene glycol powder, dialysed against 10 volumes of serum-free medium and assayed for optimal concentration as described elsewhere¹⁵. When the cells reached a density of 8×10^5 – 1×10^6 cells ml⁻¹ they were re-seeded at 1×10^5 cells ml⁻¹. On day 7 after the onset of MLC (before growth in TCGF) and 34 days after the onset of MLC (after 27 days growth in TCGF) the effector cells were washed extensively and tested for their ability to lyse ⁵¹Cr-labelled autologous LCL cells and normal lymphocytes in 6 hr ⁵¹Cr release assays as previously described⁴. The % specific ⁵¹Cr released from target cells was calculated as follows

$$\frac{\text{c.p.m. test release} - \text{c.p.m. spontaneous release}}{\text{c.p.m. maximal release} - \text{c.p.m. spontaneous release}} \times 100$$

where spontaneous release = c.p.m. released from target cells in medium alone and maximal release = c.p.m. released from target cells in detergent.

CTLs 7 days after pool sensitisation. Non-sensitised T cells from individuals J and H, following 27 days growth in TCGF, mediated low levels of cytotoxicity against autologous LCL cells.

Similar experiments to propagate CTLs cytotoxic for autologous leukaemia cells (Fig. 1) indicate that 7 days after pool sensitisation, T cells from a patient with hairy cell leukaemia mediated 27% specific ⁵¹Cr release from the autologous leukaemia cells; after 26 days growth in TCGF 80% lysis of autologous leukaemia cells occurred whereas the patient's non-sensitised T cells grown in TCGF for the same length of time caused only 9.5% specific ⁵¹Cr release. Normal autologous lymphocytes were not lysed by pool-sensitised T cells before or after growth in TCGF. In this and other experiments, CTLs grown for 3–5 weeks in TCGF were 3–4.5-fold more cytotoxic for autologous leukaemia cells than were CTLs 7 days after pool sensitisation. These findings are consistent with previous reports that mouse CTLs directed against syngeneic leukaemia cells are more cytotoxic after growth in TCGF^{12,15}.

With some leukaemia patients we have failed to detect cytotoxicity against autologous leukaemia cells after pool sensitisation, or after sensitisation in three-cell protocols with autologous leukaemia cells and allogeneic normal cells. In at least some of these cases, it is possible that CTLs are generated against the autologous leukaemia cells but that the level of cytotoxicity is below that detectable by our assay. We thus investigated whether growth of 'sensitised' cells in TCGF in such cases would result in detection of anti-leukaemic CTLs, as sensitised cells grown in TCGF are more cytotoxic than CTLs 7 days after sensitisation. T cells from a patient in relapse with acute myelogenous leukaemia mediated no cytotoxicity against autologous leukaemia cells on day 7 after sensitisation with the pool or after sensitisation in the three-cell protocol (Table 2a).

In contrast, after the sensitised cells had been grown for 11 days in TCGF, the effector cells were highly cytotoxic. When, later, with the patient in remission, T cells were isolated and subjected to pool sensitisation or sensitisation in the three-cell protocol, once again no cytotoxicity was detected on the leukaemia cells (Table 2b). However, after 19 days of growth in TCGF, high levels of cytotoxicity were detected against autologous leukaemia cells but not against autologous remission bone marrow cells which were, however, lysed by pool-sensitised T cells from unrelated individual N.

That lysis of autologous LCL and leukaemia cells by sensitised lymphocytes following growth in TCGF is mediated by CTLs rather than allo- or TCGF-activated natural killer (NK) cells is supported by the following findings. First, 92–95% of the effector cells grown in TCGF formed rosettes with sheep red blood cells (SRBC). Second, NK cells express receptors for the Fc portion of IgG and can be removed by adsorption to SRBC monolayers coated with anti-SRBC IgG^{17,18}. Nonsensitised and pool-sensitised human lymphocytes cultured in fetal calf serum are markedly cytotoxic for NK-sensitive target cells and most of this cytotoxic activity can be removed by depletion of Fc receptor-bearing cells (ref. 19 and J.M.Z., unpublished observations). In contrast, we have found that depletion of Fc receptor-bearing cells from pool-sensitised cells generated in human serum does not reduce cytotoxicity for allogeneic NK-sensitive targets or autologous abnormal cells¹⁸. Third, we wondered whether leukaemia cells would be lysed by autologous NK cells activated by interferon. Although interferon treatment augments NK-cell activity against allogeneic NK-sensitive leukaemia cells^{18,20}, cells treated with interferon were not cytotoxic for autologous leukaemia cells (J.M.Z., unpublished observations), indicating that even if TCGF can activate NK cells, the

Table 2 Lysis of autologous acute myelogenous leukaemia cells by *in vitro* sensitised T cells after growth in TCGF

Responding T cells	Stimulating cells	Days after sensitisation	Days grown in TCGF	Effector: target ratio	% Specific ^{51}Cr release $\pm$ s.d.	
					Leukaemia cells	Remission marrow cells
<i>a</i> P (patient)	L _x (Patient's leukaemia cells)	7	0	40:1	-2.1 $\pm$ 3.3	
	+ A _x (normal allogeneic lymphocytes)	18	11	40:1	41.3 $\pm$ 8.3	
				10:1	15.5 $\pm$ 6.6	
P	pool _x	7	0	40:1	-2.0 $\pm$ 3.5	
		18	11	40:1	47.7 $\pm$ 5.3	
				10:1	12.4 $\pm$ 3.8	
N (normal individual)	pool _x	7	0	40:1	47.5 $\pm$ 4.2	
<i>b</i> P (patient)	L _x (patient's leukaemia cells)	7	0	30:1	-7.2 $\pm$ 6.5	1.2 $\pm$ 2.4
	+ B _x (normal allogeneic lymphocytes)	26	19	30:1	19.2 $\pm$ 6.5	3.2 $\pm$ 5.6
P	pool _x	7	0	30:1	-2.5 $\pm$ 4.9	-6.4 $\pm$ 5.4
		26	19	30:1	30.2 $\pm$ 9.8	4.0 $\pm$ 3.6
N (normal individual)	pool _x	7	0	30:1	38.0 $\pm$ 4.7	44.6 $\pm$ 6.1

T cells were isolated as described in Table 1 from the peripheral blood of patient P who was in relapse (*a*) and remission (*b*) from acute myelogenous leukaemia. T cells (5×10^6) were cultured with 5×10^5 X-ray irradiated (4,000 R) autologous leukaemia cells (designated L_x) and 5×10^5 X-ray irradiated (2,500 R) allogeneic normal lymphocytes (designated A_x) or with 5×10^5 X-ray irradiated (2,500 R) allogeneic normal lymphocytes pooled from 20 unrelated individuals (designated pool_x) as described in Table 1. On day 7 after the onset of MLCs, the sensitised cells were collected and seeded in medium containing 10% TCGF, as described in Table 1. On day 7 after sensitisation and following growth in TCGF, the cells were tested for their ability to lyse ^{51}Cr -labelled autologous leukaemia cells and remission bone marrow cells in 6-hr ^{51}Cr -release assays as detailed in Table 1. The leukaemia cells and remission bone marrow cells, which had previously been control-rate frozen and stored in liquid nitrogen, were thawed and cultured in medium for 12 h before use as target cells. 92–94% of the patient's sensitized cells, after growth in TCGF, formed rosettes with SRBC.

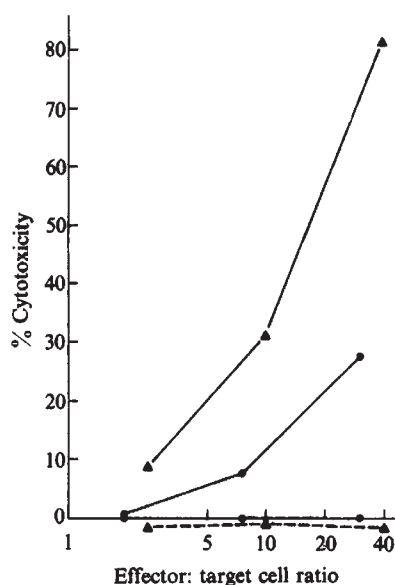


Fig. 1 Lysis of autologous hairy cell leukaemia cells (—) and autologous normal T lymphocytes (---) by T cells sensitised for 7 days with X-ray irradiated pooled allogeneic normal cells before growth in TCGF (●) and after 26 days growth in TCGF (▲). The T cells were isolated, sensitised, grown in TCGF and assayed for cytotoxicity as described in Table 1. The leukaemia cells were isolated from the patient's spleen⁵ and were control-rate frozen and stored in liquid nitrogen until 12 h before use as target cells, at which time they were thawed and cultured in medium. 93% of the pool-sensitized cells grown in TCGF formed rosettes with SRBC. Non-sensitized T cells, grown in TCGF for 26 days, mediated only 9.5% ^{51}Cr release from autologous leukaemia cells at an effector: target cell ratio of 40:1.

leukaemia cells do not seem to be susceptible to lysis by activated autologous NK cells.

The results presented here may be important for the following reasons. First, past studies aimed at generating CTLs *in vitro* against autologous or HLA-identical leukaemia cells have demonstrated that in some cases no cytotoxicity is detected following sensitisation^{2,3,5}. Our finding that such sensitised T cells cultured in TCGF exhibit significant cytotoxicity for autologous leukaemia cells suggests that the leukaemia cells of at least some of these patients express target antigens against which autologous CTLs can be generated. This approach may also be useful in other systems in which CTLs may be stimulated but in numbers that cannot be detected before growth in TCGF. Second, for basic studies and possible immunotherapeutic application, it is desirable to have large numbers of T lymphocytes highly cytotoxic for autologous leukaemia cells. Growth of *in vitro* generated anti-leukaemic CTLs in TCGF would be a useful approach to this problem, as the number of CTLs can be increased at least 5×10^6 times by 3 weeks growth in TCGF. As continuous cultures of CTLs cytotoxic for syngeneic mouse leukaemia cells inhibit leukaemia development *in vivo*²¹, it is possible that continuous cultures of human CTLs cytotoxic for autologous leukaemia cells may be of immunotherapeutic value in leukaemia patients.

We thank L. Eskra and R. Klopp for technical assistance and M. Howard and C. Smith for help in preparation of this manuscript. This work was supported by NIH grants CA-20409, CA-26836 and AI-11576 and ACS grant IM 200. J.M.Z. is the recipient of a Scholar Award from the Leukaemia Society of America.

JOYCE M. ZARLING
FRITZ H. BACH

Immunobiology Research Center and
Departments of Human Oncology,
Surgery and Medical Genetics,
University of Wisconsin,
Madison, Wisconsin 53706

Received 13 June; accepted 16 July 1979.

- Zarling, J. M., Raich, P. C., McKeough, M. & Bach, F. H. *Nature* **262**, 691–693 (1976).
- Lee, S. K. & Oliver, R. T. D. *J. exp. Med.* **147**, 912–922 (1978).
- Sondel, P. M., O'Brien, C., Porter, L., Schlossman, S. F. & Chess, L. *J. Immun.* **117**, 2197–2203 (1976).
- Zarling, J. M. & Bach, F. H. *J. exp. Med.* **147**, 1334–1340 (1978).
- Zarling, J. M., Robins, H. I., Raich, P. C., Bach, F. H. & Bach, M. L. *Nature* **274**, 269–271 (1978).
- Rollinghoff, M. & Wagner, J. *J. natn. Cancer Inst.* **51**, 1317–1318 (1973).
- Treves, S. J., Cohen, I. R. & Feldman, M. *J. natn. Cancer Inst.* **54**, 777–780 (1975).
- Bernstein, I. D. *J. Immun.* **118**, 122–127 (1977).
- Cheever, M. A., Kempf, R. A. & Fefer, A. *J. Immun.* **119**, 714–718 (1977).
- Morgan, D. A., Ruscetti, F. W. & Gallo, R. C. *Science* **193**, 1007–1008 (1976).
- Bonnard, G. D. *et al.* in *Human Lymphocyte Differentiation. Its Application to Human Cancer* (ed. Serrou, B.) (North-Holland, Amsterdam, in the press).
- Gillis, S. & Smith, K. A. *Nature* **268**, 154–156 (1977).
- Gillis, S., Baker, P. E., Ruscetti, F. W. & Smith, K. A. *J. exp. Med.* **148**, 1093–1098 (1978).
- Strausser, J. L. & Rosenberg, S. A. *J. Immun.* **121**, 1491–1495 (1978).
- Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120**, 2027–2032 (1978).
- Bentwich, Z., Douglas, S. D., Skutelsky, E. & Kunkel, H. G. *J. exp. Med.* **137**, 1532–1537 (1973).
- West, W. H., Cannon, G. B., Kay, H. D., Bonnard, G. D. & Herberman, R. B. *J. Immun.* **118**, 355–361 (1977).
- Zarling, J. M., Eskra, L., Borden, E., Horoszewicz, J. & Carter, W. J. *Immunology* (in the press).
- Ortaldo, J. R., Bonnard, G. D. & Herberman, R. B. *J. Immun.* **119**, 1351–1357 (1977).
- Herberman, R. B., Ortaldo, J. & Bonnard, G. D. *Nature* **277**, 221–223 (1979).
- Smith, K. A., Gillis, S. & Baker, P. E. *Proc. Am. Ass. Cancer Res.* **20**, 93 (1979).

Thy-1 antigen on astrocytes in long-term cultures of rat central nervous system

THE THY-1 ANTIGEN has been exploited by immunologists as a marker for thymus-derived lymphocytes in mice¹. Since its first description by Reif and Allen, it has been known to be present in equivalent amounts in brain and thymus². However, in contrast to thymus, where adult levels of Thy-1 are present at birth, Thy-1 in mouse and rat brain increases with physiological maturation of the brain from 1–2% at birth to 100% of the adult level by 3–6 weeks of age^{2–5}. It is widely assumed that neurones are the principal cell type which express Thy-1 and by implication that the increase in brain Thy-1 during neonatal development is related to increases in neuronal membrane as a result of dendritic arborisation⁵. I describe here the use of monoclonal anti-Thy-1 antiserum in combination with other recently defined neural cell markers⁶ to study the distribution of Thy-1 on cells grown in long-term cultures of rat central nervous system. I have found that Thy-1 appears on astrocytes as well as neurones and fibroblasts but not on oligodendrocytes, and that the rate of appearance of Thy-1 on cultured astrocytes parallels that of Thy-1 in brain *in vivo*. This is the first report to identify unambiguously Thy-1 on astrocytes. It settles a controversial point about the Thy-1 distribution on neural cells and clearly demonstrates that previous implications (see discussion) that neurones are the major Thy-1-positive cells in the brain is incorrect.

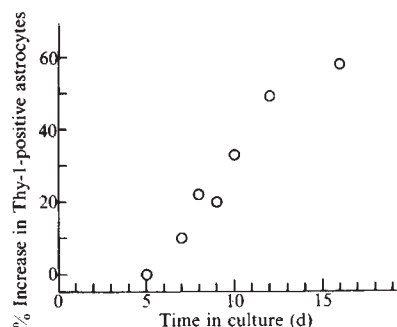


Fig. 1 Increase in percentage of Thy-1-positive astrocytes with time in culture. Cells from 5–6-d old rat corpus callosum were grown and studied by indirect immunofluorescence⁶. Data were collected from three separate experiments counting at least 150 GFAP-positive cells per coverslip.

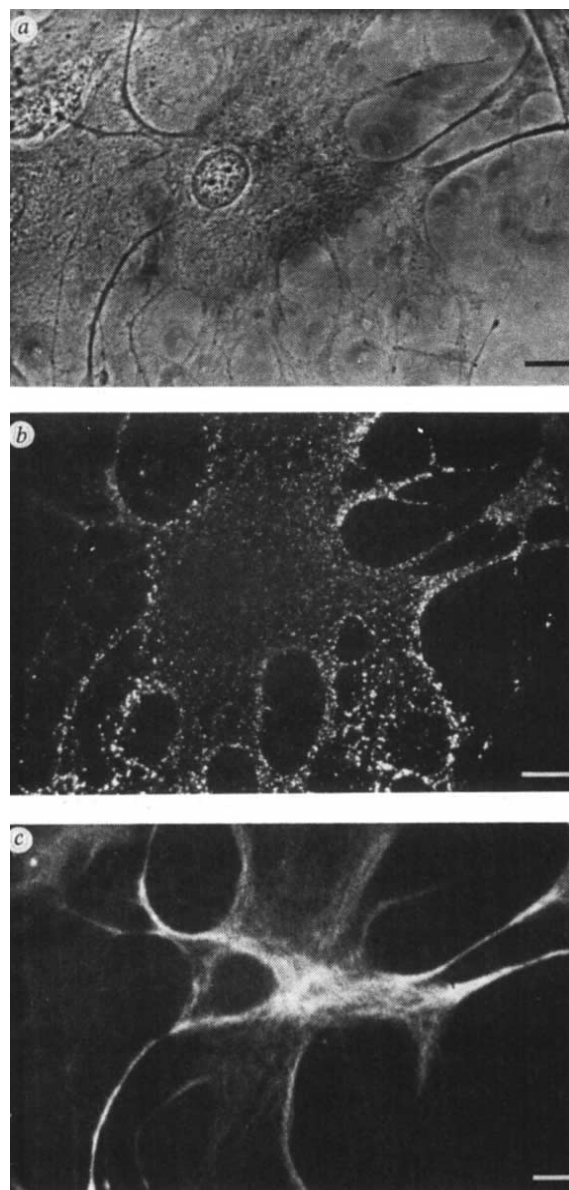


Fig. 2 Presence of Thy-1 on 'protoplasmic' astrocytes. Cells were labelled using mouse anti-Thy-1, followed by goat anti-mouse Ig-rhodamine, fixed, then labelled with rabbit anti-GFAP followed by goat anti-rabbit Ig-fluorescein⁶ and viewed with phase contrast (a); rhodamine (b), or fluorescein optics (c). Scale bar, 10 μ m.

Cells were cultured from the cerebellum and corpus callosum of 5–6-d old rats and examined by indirect immunofluorescence⁶ using the antisera listed in Table 1. In agreement with the previous report⁶, no Thy-1-positive astrocytes were seen during the first week of culture. However, by 7 d 10% of the cells positive for glial fibrillary acidic protein (GFAP), a specific marker for astrocytes, were Thy-1 positive; and after 16 d in culture slightly over half the astrocytes were Thy-1 positive (Fig. 1). Thy-1 was found on astrocytes having a range of morphologies including putative fibrous and protoplasmic⁶ astrocytes (Figs 2, 3). The possibility that fibroblasts were developing GFAP with time (that Thy-1-positive cells were becoming GFAP positive) was ruled out by the finding that none of the GFAP-positive cells in these cultures was positive for fibronectin, a marker found on the vast majority of fibroblasts in these cultures⁶. Throughout these studies, the one major central neural cell type which remained Thy-1 negative was the oligodendrocyte. Even after months in culture neither primary nor secondary corpus callosum-derived oligodendrocytes

developed Thy-1. The Schwann cell in the peripheral nervous system has also been reported to remain Thy-1 negative even after months in culture^{7,8}.

An important question remaining from the study of long-term cultures derived from neonatal rats is whether or not astrocytes become Thy-1 positive. By 3 weeks these cultures had become too confluent to count accurately and the percentage of Thy-1-positive astrocytes varied between 30 and 60%. Similar results were obtained when the monolayers were dissociated and assayed for Thy-1 and GFAP as single cell suspensions. Although the percentages of Thy-1-positive astrocytes obtained using cell suspensions tended to be somewhat lower, they are within the range of variability seen in monolayer cultures older than 16 d (30% compared with 58%). To investigate whether all astrocytes in adult brain are Thy-1 positive, cultures of cerebellum and corpus callosum were prepared from a 7-month old rat using the same method as for neonatal tissue⁶. Again, only 30–60% of the astrocytes were Thy-1 positive when the cells were examined by immunofluorescence 9 d after plating. Whether Thy-1-negative astrocytes in the adult and long-term neonatal cultures represent a small pool of immature astrocytes with an increased proliferative/survival capacity in culture

Table 1 Antisera used to identify neural cells in culture

Antiserum	Dilution	Neural cell type	Source
Mouse anti-Thy-1*	1:2,500	1 —† m	E. A. Clark & P. Lake
Rabbit anti-GFAP‡	1:100	Astrocytes ^{6,22}	—
Rabbit anti-galactocerebroside	1:40	Oligodendrocytes ^{6,22}	M. C. Raff
Rabbit anti-fibronectin	1:75	Fibroblasts ⁶	E. Ruoslahti
Goat anti-rabbit Ig-fluorescein	1:100	—	Nordic
Goat anti-mouse Ig-rhodamine§	1:60	—	Nordic

*Monoclonal IgM antiserum pooled from hybridoma tumour-bearing mice. The hybridoma was raised from spleen cells of mice immunised with rat thymocytes and screened for cytotoxicity against AKR (Thy-1.1) versus AKR/Cum or CBA (Thy-1.2) thymocytes. A description of the preparation and specificity will be described elsewhere²⁵.

†Previously found on neurones and fibroblasts^{6,17}.

‡Glial fibrillary acidic protein (GFAP) and antiserum were prepared by the method of Dahl and Bignami^{23,24}. The specificity for cultured astrocytes has been described previously^{6,22}.

§Absorbed with Sepharose-coupled rabbit immunoglobulin.

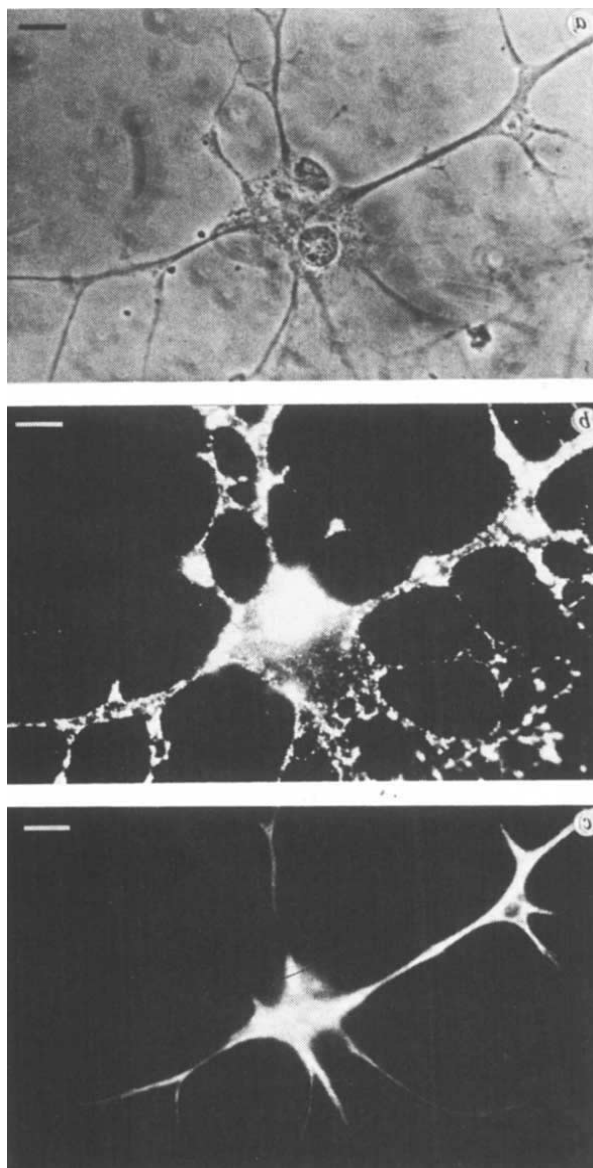


Fig. 3 Presence of Thy-1 on 'fibrous' astrocytes. Cells were labelled as described in Fig. 2 and viewed using phase contrast (a), rhodamine (b) or fluorescein optics (c). Scale bar, 10 μ m.

remains to be determined. Alternatively, the percentage of Thy-1 negative and positive astrocytes may reflect functionally distinct subclasses *in vivo*.

This is the first direct demonstration of Thy-1 antigen on glial cells. Previous attempts to establish which cells of the nervous system are Thy-1 positive have used absorption studies^{5,9–12}, and immunofluorescence on brain sections^{5,9,13} and cell lines^{14–16}. Most of these studies have suggested that Thy-1 is primarily expressed by neurones. The absorption studies have largely ignored the presence of contaminating plasma membranes from non-neuronal elements, and immunofluorescence microscopy on brain sections are technically limited in their ability to localise Thy-1 to the membranes of specific cell types within a Thy-1-positive region. Thus, although these types of studies have associated Thy-1 with neurones they could not eliminate the possibility that glia could be Thy-1 positive. Examination of neural tumour-derived cell lines has suffered from ambiguity due to the lack of criteria for classifying the various cell types^{14–16}. Previous studies using primary central nervous system-derived cultures confirmed that neurones and fibroblasts are Thy-1 positive but failed to demonstrate Thy-1 on glial cells^{6,17}. This failure was probably due to the inability to distinguish astrocytes from fibroblasts on morphological grounds¹⁷, whereas in other studies the animals and/or cultures were of insufficient age^{6,17}. Thus, the astrocyte can be added to the growing list of Thy-1-positive cell types, which in addition to lymphocytes, neurones and astrocytes include epidermal cells¹⁸, fibroblasts¹⁹, breast epithelium²⁰ and myoblasts²¹. The time course of Thy-1 development on astrocytes in culture was similar to that reported for whole neonatal rat brain^{4,5}, days 5–16 in culture corresponding to postnatal days 11–22. These findings suggest astrocytes as well as neurones contribute to both the ontological appearance and the large amount of Thy-1 present in brain.

I thank Drs Ed Clark and Phil Lake for providing the monoclonal anti-Thy-1, Dr Erkki Ruoslahti for rabbit anti-fibronectin, and Dr Martin Raff for rabbit anti-galactocerebroside as well as for helpful discussion and suggestions. This work was supported in part by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research.

REBECCA M. PRUSS

MRC Neuroimmunology Project,
Department of Zoology,
University College London,
Gower Street, London WC1, UK

Received 4 May; accepted 16 July 1979.

- Raff, M. C. *Nature* **224**, 378-379 (1969).
- Rief, A. E. & Allen, J. M. V. *J. exp. Med.* **120**, 413-435 (1964).
- Reif, A. E. & Allen, J. M. *Nature* **209**, 523 (1966).
- Douglas, T. C. *J. exp. Med.* **136**, 1054-1061 (1972).
- Barclay, A. N. *J. Neurochem.* **32**, 1249-1257 (1979).
- Raff, M. C., et al. *Brain Res.* (in the press).
- Brookes, J. P., Fields, K. L. & Raff, M. C. *Nature* **266**, 364-366 (1977).
- Fields, K. L., Brookes, J. P., Mirsky, R. & Wendon, L. M. B. *Cell* **14**, 43-51 (1978).
- Moore, M. J., Dikkes, P., Reif, A. E., Romanul, F. C. A. & Sidman, R. L. *Brain Res.* **28**, 283-293 (1971).
- Stohl, W. & Gonatas, N. K. *J. Immun.* **119**, 422-427 (1977).
- Acton, R. T., Addis, J., Carl, G. F., McClain, L. D. & Bridgers, W. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3283-3287 (1978).
- Schachner, M. *Brain Res.* **56**, 382-386 (1973).
- Barclay, A. N. & Hyden, H. J. *Neurochem.* **31**, 1375-1391 (1978).
- Schachner, M. *Nature new Biol.* **243**, 117-119 (1973).
- Acton, R. T. & Pfeiffer, S. E. *Dev. Neurosci.* **1**, 110-117 (1978).
- Lesley, J. F. & Lennon, V. A. *Brain Res.* **135**, 109-120 (1978).
- Mirsky, R. & Thompson, E. J. *Cell* **4**, 95-101 (1975).
- Scheid, M., Boyse, E. A., Carswell, E. A. & Old, L. J. *J. exp. Med.* **135**, 938-955 (1972).
- Stern, P. L. *Nature new Biol.* **246**, 76-78 (1973).
- Hilgers, J. et al. *J. natn. Cancer Inst.* **54**, 1320-1329 (1975).
- Lesley, J. F. & Lennon, V. A. *Nature* **268**, 163-165 (1977).
- Raff, M. C. et al. *Nature* **274**, 813-816 (1978).
- Dahl, D. *Biochim. biophys. Acta* **420**, 142-154 (1976).
- Dahl, D. & Bignami, A. *Brain Res.* **116**, 150-167 (1976).
- Clark, E. A. & Lake, P. *Eur. J. Immun.* (in the press).

Characteristics of *Aedes albopictus* cells persistently infected with dengue viruses

DENGUE VIRUSES, which have four serotypes, belong to the genus flavivirus of Togaviridae and are transmitted by mosquitoes such as *Aedes aegypti* or *Aedes albopictus*. Human infection with dengue viruses is manifested by disease varying in severity from mild undifferentiated fever to grave haemorrhagic fever, which is widespread in Southeast Asia¹⁻³. Persistently infected cultures can be established when cultured mosquito cells are infected with Togaviridae, either alphaviruses or flaviviruses⁴⁻⁷. Analysis of mosquito cells persistently infected with alphaviruses has shown characteristics such as the generation of small-plaque and temperature-sensitive (ts) viruses⁷⁻⁹, and resistance to superinfections with the homologous but not the heterologous viruses^{6,10}. In contrast, little is known about mosquito cells persistently infected with flaviviruses^{4,11,12}. Here, in experiments with *A. albopictus* (Singh) cells persistently infected with dengue viruses, I show that such cultures produce ts viruses and become resistant to all the types of dengue viruses.

Dengue virus type 1 (DEN-1) Hawaiian, type 2 (DEN-2) New Guinea B, type 3 (DEN-3) H-87, and type 4 (DEN-4) H-241 strains were used to prepare stock viruses¹³. Infectivities were assayed on BHK21 cells by the fluorescent focus counting method^{13,14} and were expressed as focus forming units (FFU). Replicate cultures of Singh's¹⁵ *A. albopictus*, clone C6/36 cells, were prepared in 2-ounce bottles by seeding 5×10^5 cells in 5 ml of cell growth medium (10% fetal calf serum in Eagle's medium in Earle's saline supplemented with 0.2 mM each of non-essential amino acids) in each bottle¹³. After 3 days of incubation at 28 °C, medium was removed and dengue viruses were inocu-

lated at input multiplicities (MOI) of 0.1 FFU per cell for DEN-1, DEN-3 and DEN-4 viruses, and 1 FFU per cell for DEN-2 virus. After 2 h of adsorption, cells were covered with maintenance medium (cell growth medium in which serum was reduced to 2%), and were incubated at 28 °C. Every culture showed some cytopathic effect after 2-6 days of infection. Seven days after infection, medium was replaced by fresh cell growth medium and the cultures were further incubated. When surviving cells grew out, they were transferred weekly at 1:20 dilution, and virus titres in the medium before the cell transfer were assayed on BHK21 cells at 37 °C and 28 °C. As shown in Fig. 1, the difference between the titres assayed at these two temperatures was not significant in the case of DEN-1, DEN-2 or DEN-4 infected cells up to 50 weeks of infection. However, with DEN-3 infected cells, infectivity at 37 °C was significantly lower than that at 28 °C after 22 weeks of infection, and virtually no infectivity was detected at 37 °C after 44 weeks of infection.

Medium was collected from the persistently infected cells after 50-60 weeks of infection and virus infectivities were assayed on BHK21 cells at various temperatures and by intracerebral inoculation into suckling mice (SMB). Table 1 shows that viruses obtained from persistently infected cultures were essentially temperature sensitive compared with standard viruses used to initiate the infections. Except for DEN-1, these ts viruses were not able to kill suckling mice and seemed less virulent than the standard viruses, such as found in other Togavirus systems^{4,16-18}. These viruses from persistently infected cells were neutralised to the same extent as the standard viruses by rabbit antiserum raised against the homologous standard virus.

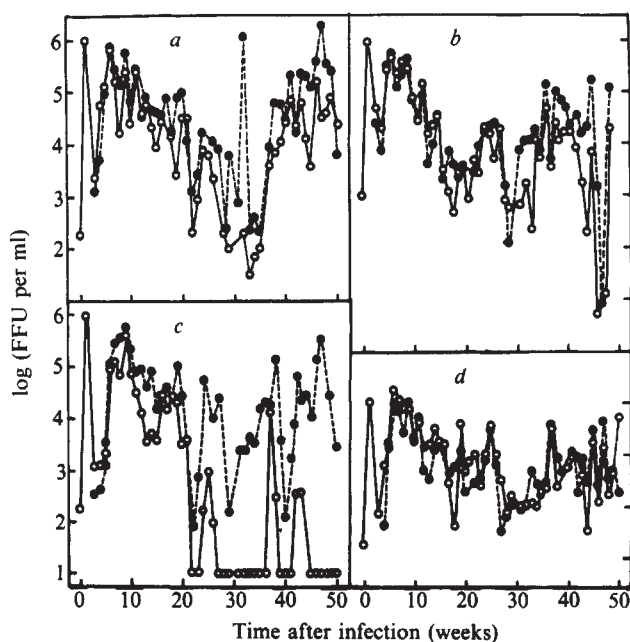


Fig. 1 Infectivity in the culture medium of *A. albopictus* cells persistently infected with dengue virus. Persistently infected cultures were established by infection with DEN-1 (a), DEN-2 (b), DEN-3 (c) and DEN-4 (d) viruses as described in the text. Virus infectivity in the medium before cell transfer was assayed at 37 °C (○—○) and at 28 °C (●-----●) on BHK21 cells.

Table 1 Infectivities of standard dengue viruses (std) and viruses produced in *A. albopictus* cells persistently infected with dengue viruses (PI)

Assay host	Temperature (°C)	DEN-1 std	DEN-1 PI	DEN-2 std	DEN-2 PI	DEN-3 std	DEN-3 PI	DEN-4 std	DEN-4 PI
BHK	39	4.9	1.0	6.7	<1.0	4.6	<1.0	6.2	2.2
	38	5.8	1.7	6.9	1.9	6.3	<1.0	6.4	3.3
	37	5.7	4.3	7.4	4.0	5.9	<1.0	6.8	4.6
	34	5.9	4.2	7.4	4.7	5.9	3.5	6.8	4.6
	28	5.8	4.8	7.0	5.2	5.5	3.7	6.6	4.9
SMB		6.0	4.0	7.1	<1.5	6.8	<1.5	6.9	<1.5

Titres are shown as log(FFU per ml) or log(LD₅₀ per ml)

Replicate cultures of *A. albopictus* C6/36 cells or the cells persistently infected with one of the four types of dengue virus were prepared in a Limbro multichamber tissue culture plate (FB-16-24-TC) by seeding 10^5 cells per ml per well. After 3 days of incubation at 28 °C, each well was inoculated with 0.1 ml of standard dengue virus as described above, or chikungunya virus

Table 2 Yields of standard dengue and chikungunya viruses from *A. albopictus* cells persistently infected (PI) with dengue viruses

Host cells	Standard virus yield				
	DEN-1	DEN-2	DEN-3	DEN-4	CHIK
C6/36	5.71±0.11	7.76±0.22	5.67±0.21	5.61±0.18	7.01±0.34
DEN-1 PI	3.98±0.27	5.91±0.26	4.34±0.39	4.11±0.50	7.11±0.45
DEN-2 PI	4.44±0.41	4.47±0.29	4.66±0.48	4.25±0.50	7.38±0.35
DEN-3 PI	3.38±0.38	5.48±0.45	2.46±0.42	2.72±0.53	6.79±0.50
DEN-r PI	3.75±0.34	5.14±0.45	3.69±0.43	3.34±0.17	6.71±0.49

Titres are shown as average of log(FFU per ml)±s.d. from 4–6 experiments.

at MOIs of 1 FFU per cell. After 2 hours of adsorption, residual virus was removed and the cells were covered with maintenance medium and incubated at 28 °C. Titres of standard viruses in the media were assayed on BHK21 cells at 38 °C after 2 days of incubation for chikungunya, 3 days for DEN-2 and DEN-4, or 4 days for DEN-1 and DEN-3 viruses, respectively.

Table 2 shows that the yields of standard dengue viruses from persistently infected cells were significantly lower than those from control C6/36 cells. The reduction of the standard virus yields was not only observed with the homologous virus used to initiate the infection, but was also seen with heterologous types of dengue viruses. However, the yield of chikungunya virus was not significantly less in the persistently infected cells than in C6/36 cells. Culture medium was collected 3 days after transfer of persistently infected cells at 60–70 weeks of infection, passed through Millipore type HA filters, and used to inoculate C6/36 cells. The cells were cultured for 1–2 weeks and then tested for the yields of standard dengue and chikungunya viruses after challenge of these viruses as described above. Yields of standard dengue viruses from those cells pretreated with medium of persistently infected cells were significantly lower than those from C6/36 cells and were similar to the yields from persistently infected cells. On the other hand, the yield of chikungunya virus was not reduced by these treatments. The dengue-resistant state in the treated cells persisted for at least 6 weeks. This transfer of dengue resistance to C6/36 cells by the culture medium from persistently infected cells did not occur when the treatment media were preheated at 56 °C for 30 min.

It is not known whether production of ts dengue viruses is unique to mosquito cells or one of the general phenomena of the persistently infected cultures¹⁹. The lower temperature of mosquito cell cultivation probably favoured the selection of ts viruses. The experimental results indicate that infection of mosquito cells with one type of ts dengue virus renders the cells resistant to all types of standard dengue viruses.

AKIRA IGARASHI

Research Institute for Microbial Diseases,
Osaka University,
Yamada-kami, Suita City,
Osaka, Japan 565

Received 16 April; accepted 12 July 1979.

- Clarke, D. H. & Casals, J. in *Viral and Rickettsial Infections of Man* (eds Horsfall, F. L. Jr & Tamm, I.) 606–658 (Lippincott, Philadelphia, 1965).
- Hammon, W. McD., Rudnick, A. & Sather, G. E. *Science* **131**, 1102–1103 (1960).
- Halstead, S. B. *Bull. Wild Hlth Org.* **35**, 3–15 (1966).
- Rehacek, J. *Acta virol.* **12**, 340–346 (1968).
- Peleg, J. *J. gen. Virol.* **5**, 463–471 (1969).
- Stollar, V. & Shenk, T. E. *J. Virol.* **11**, 592–595 (1973).
- Stollar, V., Peleg, J. & Shenk, T. E. *Immunology* **2**, 337–344 (1974).
- Shenk, T. E., Koshelnyk, K. A. & Stollar, V. *J. Virol.* **13**, 439–447 (1974).
- Igarashi, A., Koo, R. & Stollar, V. *Virology* **82**, 69–83 (1977).
- Peleg, J. & Stollar, V. *Arch. ges. Virusforsch.* **45**, 309–318 (1974).
- Banerjee, K. & Singh, K. R. P. *Ind. J. med. Res.* **56**, 812–814 (1968).
- Sinarachatanant, P. & Olson, L. C. *J. Virol.* **12**, 275–283 (1973).
- Igarashi, A. *J. gen. Virol.* **40**, 531–544 (1978).
- Igarashi, A. & Mantani, M. *Biken J.* **17**, 87–93 (1974).
- Singh, K. R. P. *Curr. Sci.* **36**, 506–508 (1967).
- Banerjee, K. & Singh, K. R. P. *Ind. J. med. Res.* **57**, 1003–1005 (1969).
- Peleg, J. *Curr. Topics Microbiol. Immun.* **55**, 155–161 (1971).
- Peleg, J. *Ann. N. Y. Acad. Sci.* **266**, 204–213 (1975).
- Preble, O. T. & Youngner, J. S. *J. infect. Dis.* **131**, 467–473 (1975).

The X-linked α -chain gene of *Drosophila* LSP-1 does not show dosage compensation

IN many species of higher organisms there is compensation for the different number of 'doses' of X-linked genes in the two sexes. This can be demonstrated by studies on enzymes which show that the level of activity of an X-linked enzyme is the same in the sex with one chromosome as in the sex with two. Dosage compensation therefore represents a system which can be used to study the control of gene activity in eukaryotes. Dosage compensation in *Drosophila melanogaster* has been extensively reviewed¹, and although it has been studied mainly in adult flies it does occur in larval stages^{2–5} and in cultured cells⁶. The one published example of an X-linked gene which does not show complete dosage compensation is the white-eosin mutant⁷ where there is more pigment in the eye of the female homozygote than in the eye of the male hemizygote. Here we report a second example, that of the X-linked α -chain gene of *Drosophila* larval serum protein-1.

Larval serum protein-1 (LSP-1)⁸ of *Drosophila* is a family of proteins made up of the random association of three polypeptides⁹ which we have called the α -, β - and γ -chains (D.B.R. and S.E.-R., in preparation). These chains are similar in their immunological properties⁹ and amino acid composition⁹ and in their peptide pattern after partial digestion with protease (H. Brock and D.B.R., unpublished results). These data suggest that they have evolved by gene duplication followed by mutation. We have mapped the coding sequences of these polypeptides (D.B.R. and S.E.-R., in preparation); the α -chain coding sequence is X-linked (39.5) whereas the other two coding sequences are found on the autosomes. Table 1 shows the amounts of α -chain (expressed as a fraction of the sum of the amounts of β - and γ -chains) found in the haemolymph of male and female larvae of two different wild-type stocks and in female larvae from a cross which generates normal diploids and larvae

Table 1 Amounts of α -chain found in the haemolymph of *Drosophila* larvae

	$\alpha/\beta + \gamma^*$		α , β and γ as % of total protein†	
	Mean	s.d.	Mean	s.d.
Oregon R females	0.42	0.04	60	4.1
Oregon R males	0.21	0.02	55.5	4.2
Samarkand females	0.51	0.06	58.1	2.6
Samarkand males	0.23	0.05	52.4	3.6
w/FM7 females	0.43	0.10	54.2	5.3
Df H368/FM7 females	0.18	0.01	52.4	5.2

Haemolymph was extracted from third instar larvae⁸ and electrophoresed on Cellologel at pH 9.3. In these conditions LSP-1 dissociates into monomers and the α -, β - and γ -chains are separated (D.B.R. and S.E.-R., in preparation). The Cellologels were stained with amido black 10B and washed thoroughly with 5% acetic acid/47.5% methanol before the three LSP-1 bands were cut out separately. The Cellologel containing each band was dissolved in 80% acetic acid and the optical density of the solution was read on a spectrophotometer at 600 nm.

* The optical density of the α -chain is expressed as a fraction of the sum of the β - and γ -chain optical densities.

† LSP-1 α -, β - and γ -chains expressed as a percentage of the total protein stained on the cellologel. Male and female larvae were sexed according to Bodenstein¹¹. The deficiency heterozygotes were obtained by crossing flies heterozygous for the deficiency HF 368 (11A2-11B9, G. Lefevre, personal communication) and the balancer chromosome FM7 (ref. 12) with male flies hemizygous for an X-chromosome marked with white (w). The two classes of female larvae were distinguished by the white malpighian tubules of the w/FM7 females compared with the normally pigmented malpighian tubules of the deficiency HF 368/FM7 larvae.

which are heterozygous for a deficiency of the α -chain gene. The results show that the LSP-1 α -chain gene shows no dosage compensation. Where there is one copy of the gene, whether it is in male larvae or in female larvae heterozygous for the deficiency, there is 50% of the quantity of α -chain found in normal females.

A possible explanation for this absence of dosage compensation is that the α -chain gene has only 'recently' arrived on the X-chromosome, either by classical translocation or as part of the process of duplicating the ancestral gene, and that it has not acquired the regulation for dosage compensation. An experimentally translocated autosomal segment on the X-chromosome does not show dosage compensation¹⁰. The α -, β - and γ -chains of LSP-1 apparently perform essentially similar functions. As they are synthesised in approximately equimolar amounts (Table 1) the ratio between the amount of these polypeptides in male and female larvae will be 5:6. This will result in less selection pressure for compensation than for a gene where the ratio of products in males and females is 1:2.

This work was supported in part by a grant from the SRC.

DAVID B. ROBERTS

SUSAN EVANS-ROBERTS

Genetics Laboratory, Department of Biochemistry,
South Parks Road, Oxford, UK

Received 14 May; accepted 24 July 1979.

1. Lucchesi, J. C. *A. Rev. Genet.* **7**, 225-237 (1973).
2. Seeoof, R. L., Kaplan, W. D. & Futch, D. G. *Proc. natn. Acad. Sci. U.S.A.* **62**, 528-535 (1969).
3. Mukherjee, A. S. & Beerman, W. *Nature* **207**, 785-786 (1965).
4. Himquist, G. *Chromosoma* **36**, 413-452 (1972).
5. Korge, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4550-4554 (1975).
6. Gvozdev, V. A., Birstein, V. Y., Kakpakov, V. T. & Polukorova, L. G. *Ontogeny* **2**, 304-310 (1971).
7. Muller, H. J. *Proc. 6th Int. Congr. Genet.* **1**, 233-235 (1932).
8. Roberts, D. B., Wolfe, J. & Akam, M. E. *J. Insect Physiol.* **23**, 871-878 (1977).
9. Wolfe, J., Akam, M. E. & Roberts, D. B. *Eur. J. Biochem.* **79**, 47-53 (1977).
10. Roehrdanz, R. L., Kitchens, J. M. & Lucchesi, J. C. *Genetics* **85**, 489-496 (1977).
11. Bodenstein, D. in *Biology of Drosophila* (ed. Demerec, M.) Ch. 4 (Wiley, New York, 1950).
12. *Drosophila Information Service* **53**, 2 (1978).

Cellular protein synthesis and inhibition of cell division are independent of butyrate-induced histone hyperacetylation

SODIUM BUTYRATE has many effects on mammalian cells—inducing the biosynthesis of new proteins¹⁻⁶, changing the relative synthetic rates of specific proteins⁷⁻¹⁰, increasing various enzyme activities^{7,9-13}, inducing morphologic changes in cultured cells^{9,12,14-19}, and inhibiting DNA synthesis and cell division^{14,15,20}. The mechanism of these effects is not known. Sodium butyrate induces massive hyperacetylation of histone in cultured mammalian cells²¹ via inhibition of histone deacetylase²²⁻²⁴; this hyperacetylation can be mimicked to varying degrees by other salts of short-chain fatty acids²². As histones interact so intimately with DNA, and as the change in extent of histone acetylation is so great, we sought to determine whether the changes in histone modification are involved in causing the changes in synthesis of specific proteins or the cessation of DNA synthesis and cell division. In hepatoma tissue culture (HTC) cells, the increased amount of histone acetylation is maximal 6-12 h after administration of sodium butyrate²², and we have used two-dimensional gel techniques²⁵ to establish whether different proteins are made during this period. Our results suggest that cellular protein synthesis and inhibition of cell division are independent of butyrate-induced hyperacetylation.

Control and butyrate-treated HTC cells were labelled with ³⁵S-methionine during 2-h pulses at the following times after butyrate administration: 0, 3, 6, 9, 12 and 24 hours. The cells were collected by centrifugation, dispersed in a lysis buffer containing 9 M urea, 5% mercaptoethanol, and 2% nonionic detergent and analysed by two dimensional electrophoresis²⁵. The protein patterns of zero time and 24-h control cells are

shown in Fig. 1a, b and that of 24-h butyrate-treated cells in Fig. 1c. Although the growth rate of control cells is significantly slowed after 24 h as they increase in density and approach stationary phase, the autoradiograms of labelled proteins from the control culture at these two times are almost identical in terms of the number and position of proteins and in terms of relative intensity of the spots.

Comparison of panels a and b of Fig. 1 with panel c shows that almost all of the spots are present in both systems, but that a significant number of proteins have changed their relative rate of synthesis (or breakdown). Proteins which are presently only in butyrate-treated cells are indicated by arrows in panel c. Only six proteins which are synthesised in the presence of butyrate cannot be distinguished in the control a and b after the period of autoradiographic exposure used. Proteins which are relatively more intense in untreated cells are indicated by circular lines in panels a and b, and proteins more intense after butyrate treat-

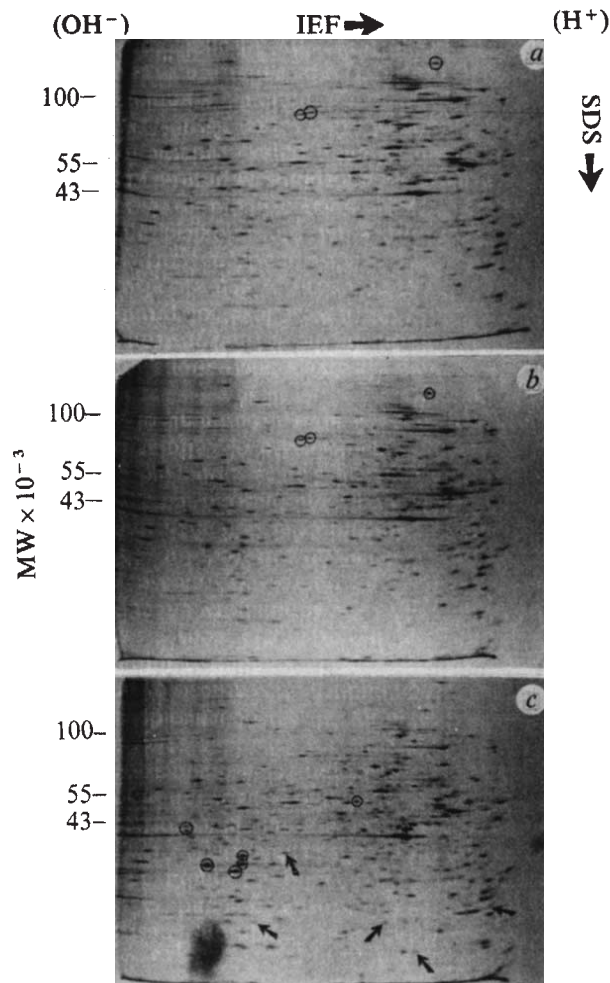


Fig. 1 Proteins synthesised in HTC cells in the presence of sodium butyrate. HTC cells, maintained in suspension culture, were incubated in Swins-77 medium in the presence of ³⁵S-methionine (100 μ Ci, 740 mCi mmol⁻¹, NEN) for 2.5 h. The cells were collected by centrifugation, washed twice in phosphate-buffered saline and disrupted in a buffer containing 9.2 M urea, 5% (w/w) 2-mercaptoethanol, 2% NP-40, 1.6% pH 5-7 ampholines and 0.4% pH 3-10 ampholines²⁵. Samples were analysed simultaneously on two-dimensional gels as described by O'Farrell²⁵. Equivalent amounts of trichloroacetic acid-precipitable c.p.m. of ³⁵S-methionine were loaded on to the gels with up to 30 μ g of total cell protein loaded per gel. The gels were dried and autoradiographed for 7 days. Panels a and b represent proteins synthesised in log phase HTC cells at zero time (300,000 cells per ml) and 24 h later. Panel c is the same as panel b except that 6 mM sodium butyrate was present during the 24 h before labelling. Spots indicated by an arrow are present only in that panel. Spots indicated by circles are present in all panels but are more intense in the panel where circled.

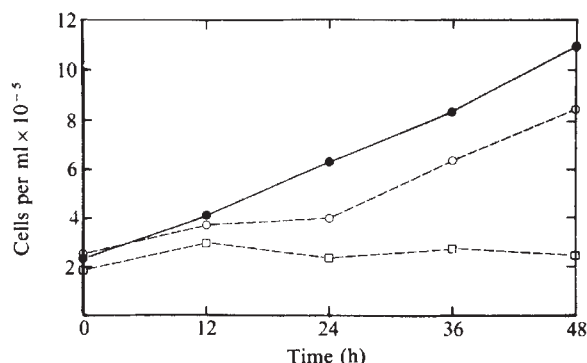


Fig. 2 Effect of salts of short chain fatty acids on cell replication. HTC cells were suspended in Swins-77 medium at a cell density of approximately 2×10^5 cells per ml in the presence of either 6 mM sodium butyrate, (□); 6 mM sodium propionate (○), or no added salts, (●). The cells were counted in a haemocytometer at the times indicated. Each point is the mean of eight determinations.

ment are shown similarly in panel c. Seven protein spots are definitely more intense in panel c relative to both a and b. This type of analysis was performed four times and only those differences which were consistently observed have been reported. Thus, out of approximately 500 proteins visualised in both control and butyrate-treated HTC cells, we have observed significant variations in the rate of synthesis or breakdown for only 16 proteins.

The concentration of butyrate administered to the cells inhibits cell growth²², so the minor changes observed could reflect a cellular adjustment to the inhibition of DNA replication rather than response to butyrate treatment. However, comparison of 24-h butyrate-treated cells with control stationary (density inhibited) cells showed that non-growing cells appeared very similar overall to both control logarithmically growing cells and butyrate-treated cells (data not shown). Some characteristic differences were observed but they were not the same variations seen between the experimental groups analysed above.

Thus, we see that for 12–18 h after the attainment of the full extent of histone acetylation there are only minor shifts in the kinds and steady state levels of proteins synthesised, at least within the pH 5–7 isoelectric focusing range. A similar analysis has been conducted at the basic end of the protein pI scale using pH 6–10 ampholines²⁶ with results that were essentially the same and accordingly such gel patterns have not been shown as they simply confirm the type of results shown above in the lower pI range. Thus there is apparently no immediate cause and effect relationship between hyperacetylation and major changes in gene expression at the translational level.

In order to find whether the hyperacetylation of histones is related to the inhibition of cell division by sodium butyrate we used other agents which affect levels of acetylation and studied their effect on cell division. As previously reported, other short chain fatty acid salts also induce hyperacetylation to almost the same extent as sodium butyrate²². However, the effect of sodium propionate induced hyperacetylation on cell division is radically different from that of sodium butyrate (Fig. 2). Although 6 mM sodium propionate causes a decrease in rate of cell division it simply amounts to a decrease of generation time from 16 to 20–22 h. In contrast, 6 mM sodium butyrate inhibits cell division after 10–20 h. Subsequently cells exposed to this level of butyrate begin to admit vital stain and some cell death ensues.

We conclude that bulk hyperacetylation of the non-H1 histones *per se* does not lead to cessation of DNA synthesis and cell division. Evidently these processes continue quite efficiently even when 70% of the H3 and H4 histones are acetylated (the level of modification induced by sodium propionate). However, as the proportion of acetylated H4 histone is consistently slightly higher (80%) after treatment with sodium butyrate^{21,22}, and in particular the level of tetra-acetylated H4 is somewhat greater, we cannot exclude the possibility that the extra margin of

hyperacetylated histone might exert a highly deleterious effect on cell division. In this regard we note that histone synthesised in the presence of butyrate is a major contributor to the high levels of modification²², and that it might be through this fraction of the total histone complement that critical differences could be developed. Finally, we note that propionate causes no major change in types of proteins synthesised as evidently by a two dimensional analysis of the type shown in Fig. 1 (data not shown).

Perhaps the most surprising thing about the recent observations on hyperacetylation of histones is not so much the dramatically increased amounts of the hyperacetylated forms of histone, but rather the apparent minimal effect on chromatin function and replication as detected by the apparent health of the cells. In view of the rapid rates of modification and hydrolysis of the monoacetyl histones in untreated cells it seems that this is likely to be the most critical modification step and that in untreated cells the small amount of the more highly modified pre-existing histones are inconsequential results of low-level recognition by the acetylating enzymes. They apparently exert no overtly deleterious effect as such, and this may account for the results described in this paper.

We thank T. Slattery for assistance in cell culture. This work was supported by NIH grants CA10871 and GM 24702.

PETER RUBENSTEIN
LINDA SEALY
STEVEN MARSHALL
ROGER CHALKLEY

Department of Biochemistry,
The University of Iowa,
Iowa City, Iowa 52242

Received 7 March; accepted 11 June 1979.

1. Fishman, P. H., Simmons, J. L., Brady, R. O. & Freese, E. *Biochem. biophys. Res. Commun.* **59**, 292 (1974).
2. Leder, A. & Leder, P. *Cell* **5**, 319 (1975).
3. Simmons, J. L., Fishman, P. H., Freese, E. & Brady, R. O. *J. Cell Biol.* **66**, 414 (1975).
4. Fishman, P. H., Bradley, R. M. & Henneberry, R. C. *Archs Biochem. Biophys.* **172**, 618 (1976).
5. Ghosh, N. K., Rukenstein, A. & Cox, R. P. *Biochem. J.* **166**, 265 (1977).
6. Ghosh, N. K. & Cox, R. P. *Nature* **267**, 435 (1977).
7. Prasad, K. N. & Sinha, P. K. *In Vitro* **12**, 125 (1976).
8. Ghosh, N. K. & Cox, R. P. *Nature* **259**, 416 (1976).
9. Tallman, J. F., Smith, C. C. & Henneberry, R. C. *Proc. natn. Acad. Sci. U.S.A.*, **74**, 873 (1977).
10. Deutsch, S. I., Ghosh, N. K. & Cox, R. P. *Biochim. biophys. Acta* **499**, 382 (1977).
11. Griffen, M. H., Price, G. H., Bazzell, K. L., Cox, R. P. & Ghosh, N. K. *Archs Biochem. Biophys.* **164**, 619 (1974).
12. Deutsch, S. I., Silvers, D. N., Cox, R. P., Griffen, M. J. & Ghosh, N. K. *J. Cell Sci.* **21**, 391 (1976).
13. Schneider, F. H. *Biochem. Pharmacol.* **25**, 2309 (1976).
14. Wright, J. A. *Expl Cell Res.* **78**, 456 (1973).
15. Ginsberg, E., Solomon, D., Sreevalsan, T. & Freese, E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2457 (1973).
16. Ghosh, N. K., Deutsch, S. I., Griffen, M. J. & Cox, R. P. *J. cell. Physiol.* **86**, 663 (1975).
17. Henneberry, R. C., Fishman, P. H. & Freese, E. *Cell* **5**, 1 (1975).
18. Henneberry, R. C. & Fishman, P. H. *Expl Cell Res.* **103**, 55 (1976).
19. Altenburg, B. C., Via, D. P. & Steiner, S. H. *Expl Cell Res.* **102**, 223 (1976).
20. Hagopian, H. K., Riggs, M. G., Swartz, L. A. & Ingram, V. M. *Cell* **12**, 855 (1977).
21. Riggs, M. G., Whittaker, R. G., Neumann, J. R. & Ingram, V. M. *Nature* **268**, 462 (1977).
22. Sealy, L. & Chalkley, R. *Cell* **14**, 115 (1978).
23. Candido, E. P. M., Reeves, R. & Davie, J. R. *Cell* **14**, 105 (1978).
24. Vidali, G., Boffa, L. C., Bradbury, E. M. & Allfrey, V. G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2239 (1978).
25. O'Farrell, P. J. *J. biol. Chem.* **250**, 4007 (1975).
26. O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 863 (1977).

Micro- and macro-stabilities of globular proteins

PROTEIN stability is usually expressed in terms of the energies required for the micro- and macro-unfolding of its structure, determined from hydrogen exchange and denaturation experiments. Both these characteristics represent qualitatively different properties of a protein macromolecule which can be classified as the rigidity of the structure and the stability of the macroscopic state. In this report, the dependence of these properties on the parameters specifying protein structure is analysed.

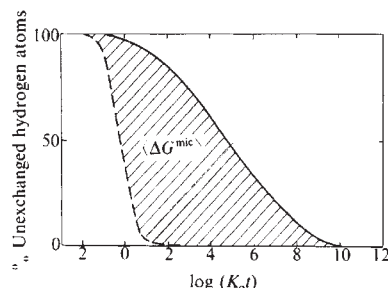


Fig. 1 Relaxation spectrum of an abstract protein (solid line) and that of a random coil polypeptide (dotted line).

There are two experimental approaches to a quantitative determination of protein stability: the first is based on the study of peptide hydrogen exchange in native protein and the other on the study of protein denaturation. The first uses the assumption that exchange of peptide hydrogen from the internal part of the protein with the hydrogen from the surrounding water takes place as a result of local or micro-unfolding of a structure¹. In this case the rate constant, K_i , of hydrogen exchange from the i th part of a macromolecule is expressed through the rate constant, K_0 , of hydrogen exchange from the unfolded conformation by

$$K_i = \exp\left(\frac{-\Delta G_i}{RT}\right) K_0 \quad (1)$$

where ΔG_i is the Gibbs energy of the corresponding micro-unfolding of the compact structure. Kinetic studies showed that different hydrogens in a protein are exchanged at different rates; the energy of unfolding of different parts of the protein varies from 0 to 14 kcal mol⁻¹ (ref. 2). The variation of the energy of micro-unfolding in proteins is usually represented by a 'relaxation spectrum' (refs 3, 4), which is a plot of the relative number of peptide hydrogens unexchanged at time t versus $\log K_0 t$ (Fig. 1). In this plot the area enclosed between the curves for the protein and the unfolded polypeptide corresponds to the average energy of micro-unfolding required for the exchange of one peptide hydrogen in the protein. This quantity, $\langle \Delta G^{\text{mic}} \rangle$, is usually considered as a measure of the micro-stability of a protein, that is, its tendency to undergo micro-unfolding.

The second approach to the determination of protein stability is based on the assumption that denaturation of a protein is a cooperative transition from the ordered native state (N) to the disordered denatured state (D), and that both these states are macroscopic, that is, they are described by thermodynamic functions⁵. In this case transition into a disordered state (denaturation) is defined by the difference thermodynamic functions of enthalpy

$$\Delta_N^D H = H^D - H^N$$

of entropy

$$\Delta_N^D S = S^D - S^N$$

and of Gibbs energy

$$\Delta_N^D G = G^D - G^N.$$

These difference functions can be determined experimentally⁶ by calorimetric measurements of the change of the enthalpy $\Delta_d H$ and the heat capacity $\Delta_d C_p$ at the temperature of denaturation T_d

$$\Delta_N^D H(T) = \Delta_d H - \int_T^{T_d} \Delta_d C_p dT \quad (2)$$

$$\Delta_N^D S(T) = \frac{\Delta_d H}{T_d} - \int_T^{T_d} \Delta_d C_p d \ln T \quad (3)$$

$$\Delta_N^D G(T) = \Delta_d H \frac{T_d - T}{T_d} - \int_T^{T_d} \Delta_d C_p dT + T \int_T^{T_d} \Delta_d C_p d \ln T \quad (4)$$

Figure 2 illustrates these functions using myoglobin as an example⁷. The $\Delta_N^D G$ function which corresponds to the work required for the transfer of a macromolecule from the native macroscopic state to the denatured macroscopic state, that is, for the disruption of all the native structure, is usually regarded as a measure of the stability of the native macroscopic state, or the macro-stability of the protein.

Because, in considering micro- and macro-stabilities, we are dealing with somewhat similar events—with the unfolding of a compact structure—one might be expected that these properties of the protein structure would be interrelated. However, in searching for the relationship between them one should not

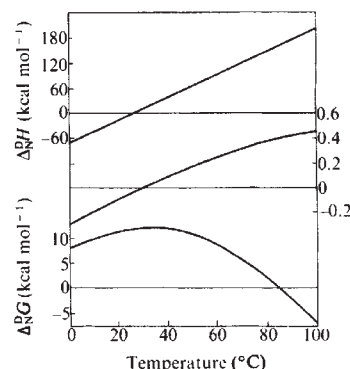


Fig. 2 Enthalpy, entropy and Gibbs energy functions for protein unfolding of myoglobin at a pH corresponding to its maximum stability⁷.

forget the great gap in the scales of these events which leads to the qualitative difference between these characteristics. It is somewhat misleading that both stabilities formally have the same dimensions, being measured in kcal mol⁻¹. However, the macrostability ($\Delta_N^D G$) is calculated per mol of the cooperative unit, whereas the micro-stability ($\langle \Delta G^{\text{mic}} \rangle$) is calculated per mol of the peptide hydrogens. As a result, the macro-stability describes the stability of the whole cooperative structure whereas the micro-stability describes the average stability of the regions of a macromolecule which are on the same scale as the micro-unfoldings. Obviously the average micro-unfolding

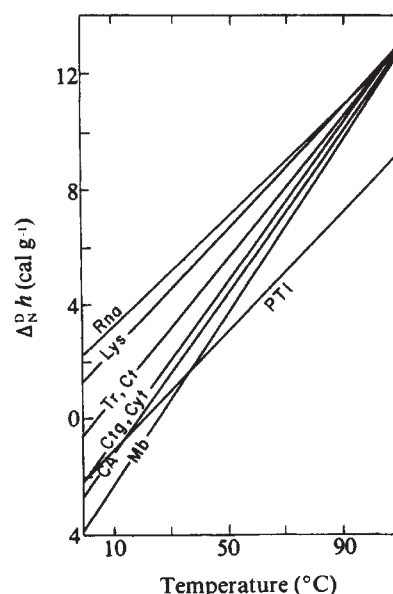


Fig. 3 Enthalpy functions for proteins listed in Table 1. The designations used are: Rna, ribonuclease A; Lys, lysozyme; Tr, β -trypsin; Ct, α -chymotrypsin; Cyt, cytochrome c ; PTI, pancreatic trypsin inhibitor; Ctg, chymotrypsinogen; CA, carbonic anhydrase B; Mb, metmyoglobin.

Table 1 Thermodynamic characteristics of some globular proteins at 25 °C

Proteins	$\langle \Delta G^{\text{mic}} \rangle$ (Kcal mol ⁻¹)	Ref.	$\Delta_N^D G$ (cal g ⁻¹)	$\Delta_N^D h$ (cal g ⁻¹)	$\Delta_N^D s (\times 10^2)$ (cal K ⁻¹ g ⁻¹)	$\Delta_d C_p$ (cal K ⁻¹ g ⁻¹)	Ref.
Ribonuclease A	2.5	8*	0.8	5.0	1.42	0.090	11
Lysozyme (egg white)	2.7	3	1.0	4.4	1.14	0.100	7
β -Trypsin	3.1	3	0.5	2.6	0.64	0.120	12*
α -Chymotrypsin	3.3	3	0.5	2.7	0.72	0.120	13
Cytochrome <i>c</i> (Ferry)	3.4	3	0.7	1.7	0.34	0.130	14
Pancreatic trypsin inhibitor	3.7	4*	0.8	0.0	-0.26	0.110	†
Chymotrypsinogen	4.7	3		2.0		0.130	15*
Carbonic anhydrase B (bovine)	6.6	3	0.4	0.1	-0.10	0.140	†
Metmyoglobin	7.1	10*	0.7	0.0	-0.22	0.150	7

$\langle \Delta G^{\text{mic}} \rangle$, averaged Gibbs energy of micro-unfolding of protein necessary to exchange one proton; $\Delta_N^D G$, specific Gibbs energy of denaturation; $\Delta_N^D h$, specific enthalpy of denaturation; $\Delta_N^D s$, specific entropy of denaturation.

* Our calculation from published data. † Unpublished results obtained in our laboratory.

cannot be large, as extended unfoldings which lead to the exchange of deeply buried hydrogens are rare. Thus, micro-stability never characterises the stability of the whole macromolecule and it can be regarded only as a characteristic of an average 'local stability' of a structure or of its 'rigidity'. It seems that this micro-stability should be treated as a specific characteristic insofar as the average size of micro-unfoldings for the exchange of one peptide hydrogen is constant for proteins. Consequently, the micro-stability can be compared only with characteristics that are also specific. Thus, we cannot compare the micro-stability directly with the macro-stability $\Delta_N^D G$, but only with the specific Gibbs energy of denaturation which is obtained by dividing $\Delta_N^D G$ by the molecular weight, M , of the cooperative unit, $(\Delta_N^D G)/M = \Delta_N^D g$. However, in using this specific Gibbs energy of denaturation we must remember that it no longer describes the stability of the macroscopic state.

Table 1 lists $\langle \Delta G^{\text{mic}} \rangle$ values and the specific thermodynamic characteristics for various proteins. We have restricted the list to globular proteins which can be considered as a single cooperative system. The reason for this restriction is that the average size of micro-unfoldings might be expected to be similar for proteins of one structural class and that the meaning of averaged quantities is more or less clear only for this simplest case. The cooperativity of all the proteins presented has been tested by the thermodynamic criterion of the coincidence of the Van't Hoff enthalpy of denaturation and the calorimetrically measured enthalpy⁷.

As seen from Table 1, the micro-stability of all nine proteins does not correlate with the specific Gibbs energy of denaturation, $\Delta_N^D g$, and does correlate with the specific enthalpy $\Delta_N^D h = (\Delta_N^D H)/M$ and with the specific entropy $\Delta_N^D s = (\Delta_N^D S)/M$: $\Delta_N^D h$ and $\Delta_N^D s$ decrease with the increase in $\langle \Delta G^{\text{mic}} \rangle$. But these correlations with the specific enthalpy and entropy seem to be in themselves the result of a correlation with the specific heat

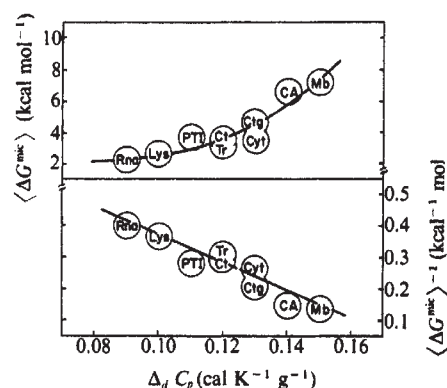


Fig. 4 Plot of the micro-stability ($\langle \Delta G^{\text{mic}} \rangle$) and the mobility ($\langle \Delta G^{\text{mic}} \rangle^{-1}$) versus the partial specific heat capacity change at denaturation ($\Delta_d C_p$) for various globular proteins. Designations as in Fig. 3.

capacity change at denaturation, $\Delta_d C_p = (\Delta_d C_p)/M$ (equations (2) and (3)).

The predominant role of $\Delta_d C_p$ in the observed correlations follows from the previously established fact that for compact globular proteins with a native structure, equally saturated with hydrogen bonds, the enthalpy and the entropy functions incline to definite points: 13 cal g⁻¹ and 3.5×10^{-2} cal K⁻¹ g⁻¹, respectively, at 110 °C (ref. 7). A consideration of the known three-dimensional structures of proteins shows (Table 2) that the concentration of hydrogen bonds, that is, the number of hydrogen bonds calculated per unit of mass $n_H = (N_H)/M$, is very similar for all studied proteins, an exception being pancreatic trypsin inhibitor, for which n_H is significantly less. Correspondingly, the denaturational enthalpy and entropy values are also lower for this protein (Fig. 3). Nevertheless, the $\Delta_d C_p$ value also correlates with the $\langle \Delta G^{\text{mic}} \rangle$ for this protein, indicating that only this correlation is of primary importance.

The relationship between $\Delta_d C_p$ and $\langle \Delta G^{\text{mic}} \rangle$ is somewhat complicated (Fig. 4). However, note that the relationship between $\Delta_d C_p$ and the inverse quantity $\langle \Delta G^{\text{mic}} \rangle^{-1}$, which can be regarded as a measure of protein structure 'mobility', seems much simpler.

It has been shown earlier⁷ that the denaturational change of the specific heat capacity of proteins is proportional to the concentration of nonpolar contacts in the native structure $n_o = (N_o)/M$, where N_o is the total number of pairs of nonpolar groups located at a distance of van der Waals contact. The validity of this statement for the proteins considered is evident from a comparison of Table 1 with Table 2. Thus, we conclude that the micro-stability or rigidity of protein structure depends on the saturation of its structure with nonpolar contacts.

Table 2 Structural characteristics of some globular proteins

Protein	M	N_H	$\frac{N_H}{M} = n_H$ ($\times 10^3$)	N_o	$\frac{N_o}{M} = n_o$ ($\times 10^3$)
Ribonuclease A	13,600	81	6.0	90	6.6
Lysozyme	14,300	89	6.2	108	7.5
α -Chymotrypsin	25,200	173	6.9	238	9.4
β -Trypsin	23,800	165	6.9	215	9.0
Cytochrome <i>c</i>	12,400	70	5.6	136	11.0
Myoglobin	17,900	133	7.4	213	11.9
Pancreatic trypsin inhibitor	6,500	28	4.3	49	7.5

M , Molecular weight; N_H , the number of internal H-bonds in the protein; N_o , the number of pairs of nonpolar groups located at a distance less than 4.0 Å in the protein.

As the parameter n_p reflects the inhomogeneity of distribution of nonpolar groups in the protein, that is, the degree of clustering of these groups, we can rephrase this conclusion by stating that the rigidity of the protein structure is determined by the clusters of nonpolar groups in the protein interior: it increases with the increase of the hydrophobic core of a macromolecule. At the same time hydrophobicity of the protein interior does not noticeably affect the stability of the protein, that is, the Gibbs energy of denaturation. The absence of this dependence can be understood if one remembers that the Gibbs energy is determined by the difference between the enthalpy and entropy terms and not by their similarity, as $\Delta G = \Delta H - T \Delta S$. As a result, n_p does not affect $\Delta_N^D G$ because it influences the enthalpy and entropy of denaturation to similar extents. But such a factor as the cross-linking of a polypeptide chain, which affects only the entropy of unfolding, can essentially shift the enthalpy-entropy balance and change $\Delta_N^D G$. Thus, the absence of any correlation between the micro-stability and the Gibbs energy of denaturation of a protein is not as surprising as it might seem.

P. L. PRIVALOV
T. N. TSALKOVA

Institute of Protein Research,
Academy of Sciences of the USSR,
Poustchino, Moscow Region, USSR

Received 12 December 1978; accepted 3 July 1979.

1. Linderström-Lang, K. H. & Schellman, J. A. *The Enzymes*, Vol. 1, 2nd edn (eds Boyer, P. D., Lardy, H. & Myrback, K.) 443–510 (Academic, New York, 1959).
2. Hvidt, A. & Nielsen, S. O. *Adv. Protein Chem.* **21**, 287–386 (1966).
3. Willumsen, L. C. *r. Trav. Lab. Carlsberg* **38**, 223–295 (1971).
4. Hvidt, A. & Walleik, K. *J. Biol. Chem.* **247**, 1530–1535 (1972).
5. Anson, M. L. *Adv. Protein Chem.* **2**, 361–368 (1945).
6. Privalov, P. L. *FEBS Lett.* **40**, S141–S152 (1974).
7. Privalov, P. L. & Khechinashvili, N. N. *J. molec. Biol.* **86**, 665–684 (1974).
8. Tiktopulo, E. I. & Privalov, P. L. *Biofizika (USSR)* **20**, 778–782 (1975).
9. Hvidt, A. & Pedersen, E. J. *Eur. J. Biochem.* **48**, 333–338 (1974).
10. Zavadzky, P., Jochansen, J. T. & Hvidt, A. *Eur. J. Biochem.* **56**, 67–72 (1975).
11. Tiktopulo, E. I. & Privalov, P. L. *Biophys. Chem.* **1**, 349–357 (1974).
12. Tischenko, V. M. & Gorodkov, V. G. *Biofizika (USSR)* **24**, 334–335 (1979).
13. Tischenko, V. M., Tiktopulo, E. I. & Privalov, P. L. *Biofizika (USSR)* **19**, 400–404 (1974).
14. Privalov, P. L. & Khechinashvili, N. N. *Biofizika (USSR)* **19**, 14–18 (1974).
15. Jackson, W. M. & Brandts, J. F. *Biochemistry* **9**, 2294–2301 (1970).

Membrane crystals of ubiquinone:cytochrome *c* reductase from *Neurospora* mitochondria

HIGH-RESOLUTION X-ray diffraction or electron microscopy studies of membrane proteins are hampered by the difficulty in obtaining crystals or periodic aggregates of the proteins. So far, only two examples of membrane proteins forming highly ordered arrays are known, the purple membrane of *Halobacterium halobium*¹ and a membranous preparation of cytochrome *c* oxidase from beef heart². Ubiquinone:cytochrome *c* reductase, a multiprotein enzyme, contributes about 10% of the inner mitochondrial membrane protein³. The enzyme is isolated as a dimer (molecular weight (MW) ~550,000) and the monomeric unit comprises two *b* cytochromes (MW ~30,000 each), a cytochrome *c*₁ (MW ~31,000), an iron-sulphur subunit (MW ~25,000) and six subunits without known prosthetic groups (MWs ~50,000, 45,000, 45,000, 14,000, 12,000 and 8,000)^{3,4}. The enzyme catalyses the reduction of ferricytochrome *c* by reduced ubiquinone and conserves the energy generated by the oxidation-reduction reaction by translocating protons across the inner mitochondrial membrane⁵. Cytochrome *c* reductase is also of interest with respect to the genetics and biogenesis of mitochondria. The cytochrome *b* subunits of this enzyme belong to the small group of proteins which are translated on mitochondrial ribosomes⁶ and coded for by mitochondrial DNA⁷. We report here a method for preparing membrane crystals of ubiquinone:cytochrome *c* reductase isolated from *Neurospora crassa*.

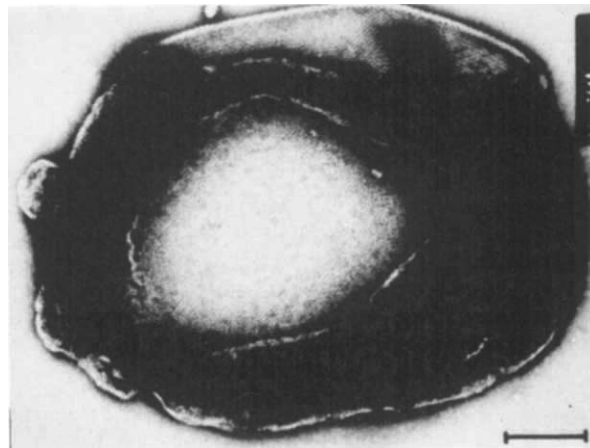


Fig. 1 Negatively stained broken vesicle of crystallised cytochrome *c* reductase. Scale bar, 0.5 μ m.

The cytochrome *c* reductase was released from the membranes of *N. crassa* mitochondria by Triton X-100. A micelle (MW ~100,000) of this non-ionic detergent replaces (most of) the phospholipids⁴. The enzyme-detergent complex was then purified by affinity chromatography on immobilised cytochrome *c* followed by gel chromatography. It has a Stokes radius of 86 Å and a sedimentation coefficient $S_{20,w}$ 15.6 (ref. 4).

To obtain membrane crystals from the isolated enzyme, the enzyme-bound detergent was replaced by phospholipid as follows. One ml of a solution containing 2.5–3 mg ml⁻¹ enzyme in 0.05% (w/v) Triton X-100, 50 mM Tris-acetate, pH 7, and 0.5 mM EDTA was mixed with 1 ml of a solution containing 4 mg ml⁻¹ phosphatidyl choline (from soybean) and 1 mg ml⁻¹ phosphatidyl serine (from bovine brain) and 0.25% Triton X-100, 50 mM Tris-acetate, pH 7, 0.5 mM EDTA and 5 μ M butylated hydroxytoluene. Triton X-100 was removed by stirring 0.5 g Bio Beads SM-2 (ref. 8) in the combined solution for 2.5 h at 4 °C. The turbid solution was then subjected to isopycnic centrifugation on a 12-ml 10–50% (w/v) sucrose-density gradient in 50 mM Tris-acetate, pH 7, 0.5 mM EDTA and 5 μ M butylated hydroxytoluene at 280,000g for 20 h. The membranes banded at a density of 1.15 and the excess phospholipid remained close to the top of the gradient. The protein to lipid ratio of the membranes was about 1.7 (w/w).

The membranes were prepared for electron microscopy as follows. The suspension was dialysed overnight against 50 mM Tris-acetate, pH 7, to remove sucrose and then brought to 0.5 mg ml⁻¹ protein. Copper grids (400 mesh) coated with collodion-reinforced carbon films were rendered hydrophilic by glow-discharge in air. Drops of the lipoprotein suspension were placed on parafilm and grids were touched to these for 25 s, to distilled water for 10 s and then to 2% aqueous uranyl acetate for 45 s. Excess liquid was removed from the grid with filter paper between each stage. The grids were examined in a Philips EM400 electron microscope operating at 80 kV and images were recorded at $\times 43,000$ –70,000 magnification. Optical diffraction was carried out as described previously⁹.

The electron micrographs showed the presence of negatively stained vesicles up to several microns in diameter which consisted of the cytochrome *c* reductase in highly ordered arrays (Fig. 1). Images of broken vesicles (Fig. 2a) showed a 'one-sided' pattern of protein packing whereas closed vesicles showed a moiré pattern (Fig. 2b) arising from overlap of the lattices from the two sides of the flattened vesicle. The overlap is clearly demonstrated by the optical diffraction pattern (Fig. 2d): the pattern from the image of the closed vesicles shows a mirror-related doubling of all the spots compared with the pattern from the one-sided image (Fig. 2c). The angle between the two lattices, as measured by the relative rotation of the spots, was always close to 90°. This restricted angle may result from a folding occurring preferentially along the line of the heavily

stained groove. Such a folding also explains the finding that the opposite edges of the vesicles are frequently straight and parallel (not shown).

The diffraction patterns extend to the 5th order, limiting the resolution to about 25 Å. This is probably due to the staining and electron beam damage, but the resolution should be extendable by low dose microscopy.

The unit cell is $137 \text{ Å} \times 174 \text{ Å}$. The diffraction pattern (Fig. 2c,d) and the computer-reconstructed image⁹ (Fig. 3) are interpreted as arising from a projection of the membrane crystal corresponding to plane group pgg. If our interpretation of the plane as being a single layer is correct, this symmetry requires that alternate protein molecules are arranged in an up-and-down manner, that is, the glide planes seen in the image correspond, in projection, to two-fold screw axes in the plane of the membrane. This is unlike the membrane crystals reported for cytochrome *c* oxidase², in which the protein is arranged unidirectionally. Furthermore, with the cytochrome *c* reductase crystals no interaction seems to occur between the sheet on opposite sides of the vesicles.

The dimensions of one dimeric cytochrome *c* reductase molecule in projection are $\sim 90 \text{ Å} \times 70 \text{ Å}$. These dimensions are only about half the Stokes diameter, 172 Å, of the (dimer) enzyme-Triton X-100 complex estimated by gel filtration⁴. Even when the amount of enzyme-bound Triton X-100—about

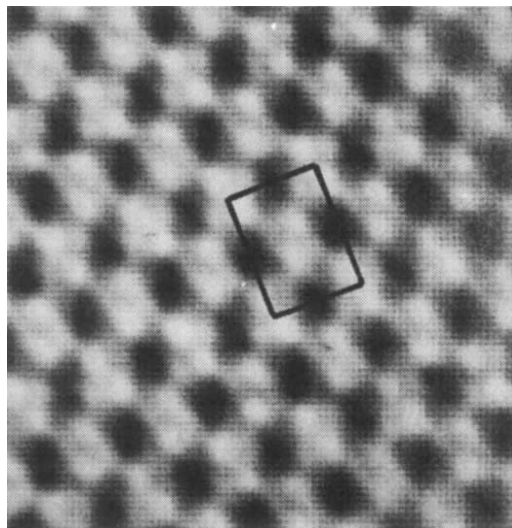


Fig. 3 Computer-reconstructed image of the 'one-sided' membrane crystal shown in Fig. 2a. The protein molecules are apparently dimeric. In projection they are arranged with pgg symmetry, so that alternate dimers are enantiomorphic and thus must arise from opposite views. The rectangular unit cell ($137 \text{ Å} \times 174 \text{ Å}$) has been drawn in.

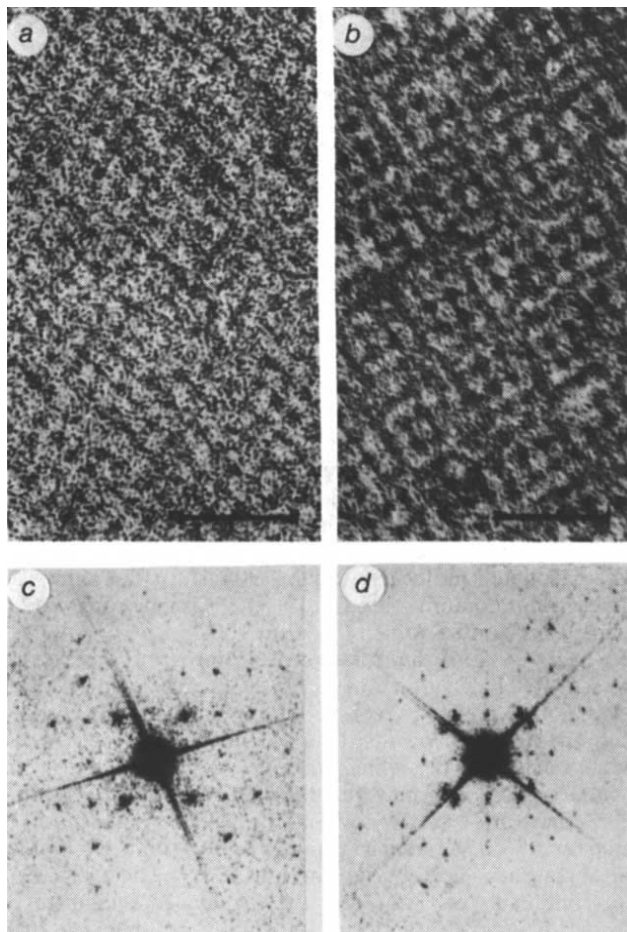


Fig. 2 a, b, Higher magnification images of the membrane crystals. A, 'one-sided' pattern obtained from the central region of the broken vesicle shown in Fig. 1. B, Moiré pattern produced by overlap of images for the two sides of a flattened closed vesicle. Scale bar, 50 nm. c, d, Optical diffraction patterns for the images of the 'one-sided' membrane crystal (C) shown in a and of the 'two-sided' membrane crystal (d) shown in b. The absence of the (0, 1) and (0, 3) reflection on the meridian in c is consistent with the pgg symmetry. In the case of d the two overlapping patterns are rotationally displaced by exactly 90°.

0.18 g per g protein—is subtracted, a Stokes diameter much larger than 100 Å would result. This inconsistency between the dimensions of the enzyme in projection and its apparent hydrodynamic diameter in solution, indicates that cytochrome *c* reductase is a rod-like enzyme which spans the membrane with its elongated axis perpendicular to the plane of the membrane.

We thank Brigitte Juchs and Annette Scharm for technical assistance, and Sir John Kendrew for his support and interest in this project.

PAUL WINGFIELD
TALMON ARAD
KEVIN LEONARD
HANNS WEISS

European Molecular Biology Laboratory,
Meyerhofstrasse 1, P. O. Box 10.2209,
6900 Heidelberg, FRG

Received 20 March; accepted 11 July 1979.

1. Unwin, P. N. T. & Henderson, R. *J. molec. Biol.* **94**, 425–440 (1975).
2. Henderson, R., Capaldi, R. A. & Leigh, J. S. *J. molec. Biol.* **112**, 631–648 (1977).
3. Rieske, J. S. *Biochim. biophys. Acta* **456**, 195–247 (1976).
4. Weiss, H. & Kolb, H. J. *Eur. J. Biochem.* (in the press).
5. Mitchell, P. *FEBS Lett.* **43**, 189–194 (1974).
6. Weiss, H. & Ziganke, B. *Eur. J. Biochem.* **41**, 63–71 (1974).
7. Tzagoloff, A., Akai, A., Needleman, R. B. & Zulch, G. *J. biol. Chem.* **250**, 8236–8242 (1975).
8. Holloway, P. S. *Analyt. Biochem.* **53**, 304–308 (1973).
9. Erickson, H. P., Voter, W. A. & Leonard, K. *Meth. Enzym.* **49**, 39–63 (1978).

Adenine *N*⁶-substituent of agrocin 84 determines its bacteriocin-like specificity

AGROCIN 84 (Fig. 1(1)) is the active factor¹ in the biological control of crown gall, a plant cancer induced by certain strains of *Agrobacterium*. Growth inhibition by agrocin 84 has been correlated with pathogenicity^{2–5} and the presence of a nopaline (*N*²-(1,3-dicarboxypropylarginine)) tumour-inducing plasmid^{3–5}. Characterisation of agrocin 84 as a disubstituted adenine nucleotide⁶ has been reported and a proposed⁷ peptide structure has been retracted⁸. What are those salient features of agrocin 84 that determine its remarkable bacteriocin-like specificity

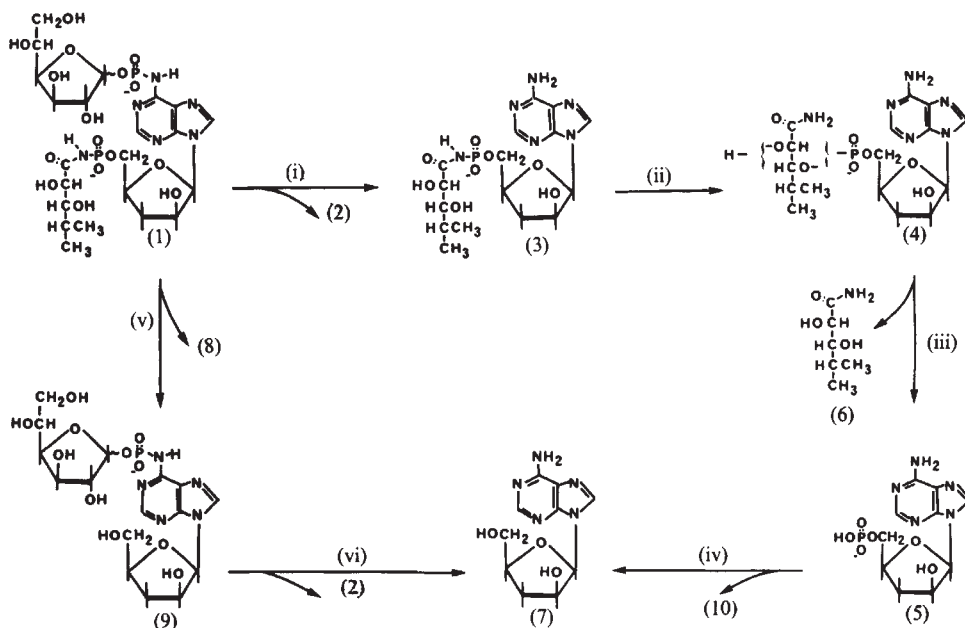


Fig. 1 Agrocin 84, acid and alkaline degradation pathways. Products: (1) agrocin 84; (2) cyclic glucofuranosyl phosphate; (3) nucleoside 5' phosphoramidate; (4) nucleoside 5' phosphodiester; (5) nucleoside 5' phosphomonoester; (6) D-threo-2,3-dihydroxy-4-methylpentanamide; (7) nucleoside (9-(3'-deoxy-β-D-threo-pentofuranosyl)-adenine); (8) cyclic phosphoramidate of D-threo-2,3-dihydroxy-4-methylpentanamide; (9) nucleoside N⁶-D-glucofuranosyloxyphosphoramidate; (10) inorganic phosphate. Reaction conditions: (i) 110°C, pH 7, 15 min; (ii) 25°C, 0.1 M HCl, 17 h; (iii) 25°C snake venom phosphodiesterase, pH 8.8, 3 h; (iv) 25°C, snake venom 5' nucleotidase, pH 9.0, 1 h; (v) 1 M NH₄OH, 25°C, 3 h; (vi) 1 M NH₄OH, 110°C, 1 h.

towards cancer-inducing strains of *Agrobacterium*? The structural evidence reported here shows that a 5'-phosphoryl linkage from the 'fraudulent' nucleoside core 9-(3'-deoxy-β-D-threo-pentofuranosyl) adenine⁶ to the amide group of D-threo-2,3-dihydroxy-4-methylpentanamide (Fig. 1(6)) is required for antibiotic activity, but bacteriocin-like specificity is conferred by a D-glucofuranosyloxyphosphoryl substituent at N⁶ of adenine. Agrocin 84 seems to be the harbinger of a series of highly specific nucleotide bacteriocins, which behave as molecular 'Trojan horses'.

The evidence leading to Fig. 1 is summarised as follows: mild alkaline hydrolysis (1 M NH₄OH at 25°C for 24 h) of agrocin 84 produces nucleotide (9) containing D-glucose, with the characteristic⁶ absorption maximum (264 nm, ε 19,500) of an adenine nucleoside N⁶-phosphoramidate. In addition, the non-UV absorbing organic phosphate monoesters from the initial cyclic product (8) yield inorganic phosphate and a vicinal glycol after dephosphorylation with *Escherichia coli* alkaline phosphatase. This glycol has now been crystallised (m.p. 192–194; [α]_D –51.6±3) and shown to be virtually indistinguishable in all its measured properties including melting point, mixed melting point, IR spectrum, mass spectrum, optical rotation, and chromatographic and electrophoretic behaviour from a synthetic sample of D-threo-2,3-dihydroxy-4-methylpentanamide (6). The racemic threo-2,3-dihydroxy-4-methylpentanoic acid (m.p. 108) described by Braun⁹ also crystallises as an isomorphous form with a broad m.p. 114–121 and was resolved by way of its quinine salt. The dextrorotatory acid ([α]_D = +13.6±0.2; m.p. 141–142) was converted by way of its methyl ester derivative to yield the required levorotatory amide (6) (m.p. 194–195; [α]_D –52.9±0.6).

Periodate oxidation of agrocin 84 (Fig. 1(1)) yields one mol of formaldehyde and one mol of isobutyraldehyde, whereas periodate oxidation of nucleotide (9) yields only one mol of formaldehyde. As no glycol constituents other than glucose were detectable in nucleotide (9) after further hydrolysis to the nucleoside and inorganic phosphate, it follows that the glucose is present as a glucofuranosyloxyphosphoramidate and that neither the 2, nor the 3 hydroxyls of the D-threo-2,3-dihydroxy-4-methylpentanamide is esterified to phosphate in agrocin 84 (1). The nature of the anomeric glucofuranosyl linkage remains uncertain. Further alkaline hydrolysis of nucleotide (9) yields the nucleoside 9-(3'-deoxy-β-D-threo-pentofuranosyl) adenine structure (7) and a cyclic glucofuranosyl phosphate (2) which in turn hydrolyses to yield glucose phosphate monoesters.

In an alternative route to the nucleoside core, brief heat treatment (100°C at pH 7 for 15 min) of agrocin 84 produces

two nucleotides (3) and (4), which do not contain glucose and exhibit absorption maxima at 260 nm. In addition, a cyclic glucofuranosyl phosphate (2) is produced. Nucleotide (3) undergoes periodate oxidation and yields one mol of isobutyraldehyde. Nucleotide (3) also forms a borate complex and is slowly cleaved by snake venom phosphodiesterase to produce the nucleoside phosphate monoester (5). Treatment of the periodate sensitive, borate complexing nucleotide (3) with dilute acid (0.1 M HCl at 25°C for 17 h) irreversibly isomerises it to the periodate insensitive, non-borate complexing nucleotide (4). The latter is readily cleaved by snake venom phosphodiesterase to produce one mol of the nucleoside phosphate monoester (5) and one mol of D-threo-2,3-dihydroxy-4-methylpentanamide (6). Finally, the nucleoside phosphate monoester (5) is smoothly cleaved by a specific snake venom 5'-nucleotidase to yield inorganic phosphate (10) and the nucleoside (7). The latter observation combined with the formation of isobutyraldehyde by the periodate oxidation of nucleotide (3) and agrocin 84 (1), suggests the presence of a second phosphoramidate linkage in agrocin 84 from the 5'-hydroxyl of the nucleoside to the amide nitrogen atom, but the possibility that the linkage to the amide group may be by way of an enolic form to the oxygen atom is not excluded by these data.

The calculated molecular weight (904.4) for the anhydrous bis-triethylammonium salt C₃₄H₆₆N₈O₁₆P₂ is in accord with the predicted⁶ maximum value of 1,100±100.

When the scale of the isolation procedures was increased to process 1,000 l of culture medium, it became evident that a second biologically active component was present. This product was soon shown to be identical to the thermal degradation product nucleotide (3), which lacks the N⁶-glucofuranosyloxyphosphoramidate and may be an isolation artefact.

For structure-activity correlations, the potency of a compound may be defined in arbitrary agrocin units per mol. An arbitrary agrocin unit is that amount of agrocin necessary to totally inhibit a zone of radius 1.0 cm (corresponding to the dashed ordinates in Fig. 2) in a lawn of the indicator strain. Using the pathogenic indicator strain 57A in the Stonier¹⁰ bioassay procedure, the logarithmic dose-response plot (Fig. 2a) shows that 20 μl 3.49×10⁻⁷ M solution agrocin 84 corresponds to one arbitrary agrocin unit; this is equivalent to 1.4±0.3×10¹¹ agrocin units per mol. Using the same indicator strain, the degradation fragment (3) shows a potency of 3.9±0.5×10⁸ units per mol. In other words, for the pathogenic strain 57A, agrocin 84 is 360 times more potent than its nucleotide fragment (3). However, Fig. 2b shows that the non-pathogenic strain 57 is not inhibited by agrocin 84 at a range of molarities at which the

nucleotide degradation fragment (3) readily demonstrates inhibition, with a measured potency of $3.2 \pm 0.5 \times 10^8$ agrocin units per mol.

Thus, it can be seen that the nucleotide fragment (3) without the N^6 -D-glucufuranosyloxyphosphoramidate substituent shows the nonspecific growth inhibition of a simple antibiotic, but only agrocin 84 (1) with its unusual N^6 substitution shows the strain specificity of a bacteriocin for the pathogenic strain 57A.

Over similar concentration ranges, neither strain was inhibited by nucleotides (4) and (9) or the nucleoside (7). The absence of detectable activity using either indicator strain with the N^6 -substituted nucleotide (9) shows that although this is a necessary condition for selectivity, it is not a sufficient requirement for antibiotic activity. For detectable antibiotic activity in the bioassay, using initial well concentrations of less than 10^{-3} M, a phosphoramidate linkage to the amphiphilic 2,3-dihydroxy-4-methylpentanamide (6) seems to be mandatory.

The pathogenic *in planta* transconjugant strain 57A was originally obtained² using the non-pathogenic strain 57 as the recipient and the pathogenic strain 27 as donor. Examination by agarose gel electrophoresis¹¹ of the isolated plasmid DNA¹² from strains 57, 57A and 27 confirmed that strain 57A differed from strain 57 solely by the presence of a band corresponding in mobility to the 1.5×10^8 , molecular weight¹³ donor strain 27 tumor-inducing plasmid (pTi27). Murphy and Roberts¹⁴ have used this subtle difference in plasmid makeup to show that the virulence plasmid (pTi27) encodes an energy-dependent transport mechanism which selectively transports intact agrocin 84 into strain 57A. The absence of the virulence plasmid (pTi27) in strain 57 leads to exclusion of agrocin 84, and it is noteworthy

that agrocin 84 does not bind irreversibly to the cell surface of susceptible strains. The analogy of a molecular 'Trojan Horse' seems apt.

A cytosine 'nucleotide bacteriocin', agrocin 108 is also under investigation (Elvin and M. E. T. unpublished) and we suggest that these cell-specific nucleotides may be more abundant in nature than is generally realised.

Support from the Australian Research grants and Charles John Everard Scholarship committees is acknowledged. We thank Dr J. K. Macleod for mass spectrometric data and Jan Nield for technical assistance.

M.E. TATE
P. J. MURPHY
W. P. ROBERTS
A. KERR

Waite Agricultural Research Institute,
University of Adelaide,
Glen Osmond, South Australia 5064

Received 23 March; accepted 28 June 1979.

1. Kerr, A. & Htay, K. *Physiol. Pl. Path.* **4**, 37-44 (1974).
2. Roberts, W. P. & Kerr, A. *Physiol. Pl. Path.* **4**, 81-92 (1974).
3. Watson, B., Currier, T. C., Gordon, M. P., Chilton, M. & Nester, E. W. *J. Bact.* **123**, 255-264 (1975).
4. Engler, G., Holsters, M., Van Montague, M., Schell, J., Hernalsteens, J. P. & Schilperoort, R. *Molec. gen. Genet.* **138**, 345-349 (1975).
5. Van Larabeke, N. *et al. Nature* **255**, 742-743 (1975).
6. Roberts, W. P., Tate, M. E. & Kerr, A. *Nature* **265**, 379-81 (1977).
7. McCardell, B. & Pootjes, C. F. *Antimicrob. Agents Chemother.* **10**, 498-502, (1976).
8. Thompson, R. J., Hamilton, R. H. & Pootjes, C. F. *Plant Physiol. Suppl.* **59** (6), 110 (1977).
9. Braun, L. *Monaish. Chem.* **17**, 207-224 (1896).
10. Stonier, T. J. *Bact.* **79**, 889-898 (1960).
11. Meyers, J. A., Sanchez, D., Elwell, L. P. & Falkow, S. *J. Bact.* **127**, 1529-1537 (1976); *J. Bact.* **129**, 1171 (1977).
12. Currier, T. C. & Nester, E. W. *Analyt. Biochem.* **66**, 431-441 (1976).
13. Sciaky, D., Montoya, A. L. & Chilton, M. D. *Plasmid* **1**, 238-253 (1978).
14. Murphy, P. J. & Roberts, W. P. *J. gen. Microbiol.* (in the press).

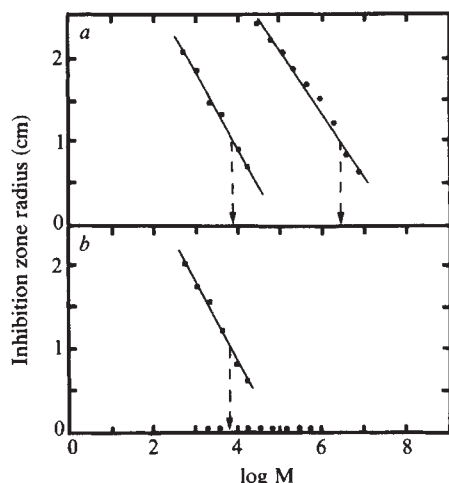


Fig. 2 Least squares fit for the negative logarithm of centre well (20 μ l) molarity (initial) against inhibition zone radius using the pathogenic *Agrobacterium radiobacter* var. *tumefaciens*, strain 57A. Comparison of agrocin 84 (●; Fig. 1(1)) with the nucleoside 5' phosphoramidate of D-threo-2,3-dihydroxy-4-methylpentanamide (■; Fig. 1(3)). *b*, As for *a*, but with the non-pathogenic *Agrobacterium radiobacter* strain 57 as the bacterial indicator lawn in the Stonier¹⁰ bioassay. Dashed ordinates indicate the negative log molarity (initial) for an arbitrary agrocin unit of potency. Note that whereas the nucleoside 5' phosphoramidate (Fig. 1(3)) shows antibiotic activity towards both strains, only agrocin 84 (Fig. 1(1)) with an N^6 -D-glucufuranosyloxyphosphoramidate substitution shows specific growth inhibition of the pathogenic strain 57A. Molar absorptivities of 19,860 at 264 nm for agrocin 84 (ref. 6) and 15,400 at 260 nm for the nucleotide 5' phosphoramidate were used for initial concentration measurements on electrophoretically and chromatographically homogeneous samples.

Erratum

In the letter 'Inhomogeneous distribution of filipin-sterol complexes in smooth muscle cell plasma membrane' by R. Montesano, *Nature* **280**, 328-329, Fig. 2 on page 329 was printed with poor contrast. The figure is reprinted below.

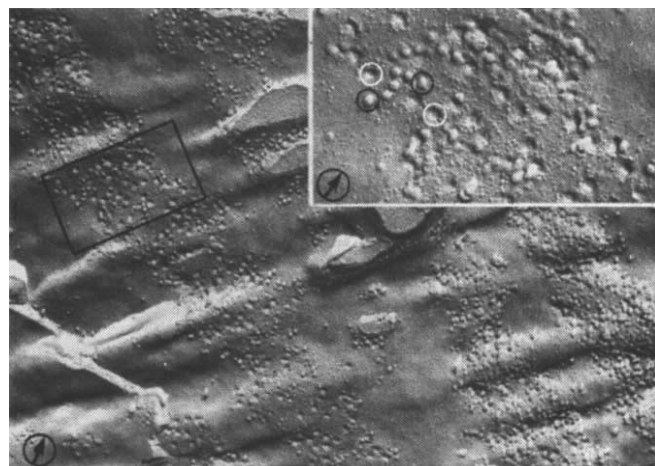


Fig. 2 Freeze-fracture replica of a smooth muscle cell fixed in glutaraldehyde-filipin mixture. The detail of a typical band of invaginations (black rectangle) is shown at high magnification in the inset. Numerous filipin-sterol complexes (black circles) are interspersed among the microinvaginations (white circles), whereas they are virtually absent outside the invaginated membrane regions. Filipin-sterol complexes appear as 25-30 nm protuberances which can be distinguished from the contiguous microinvaginations by the opposite polarity of platinum deposition (the direction of the shadowing is indicated by the encircled arrow). $\times 22,000$. Inset: $\times 44,000$.

matters arising

Dihydroergocryptine is a non-selective antagonist for human platelet α -adrenoreceptors

BINDING of ^3H -dihydroergocryptine (DHEC) which is suppressible by phen-tolamine is widely used as a measure of the α -adrenoreceptor content of tissues¹. Recently, Kunos *et al.*² have reported that phenoxybenzamine, an irreversible α_1 -selective adrenoreceptor antagonist, suppressed only a portion of the specific DHEC binding to rabbit uterine membrane fractions although this antagonist had caused total blockade of noradrenaline-induced contraction in the rabbit uterine strips from which the membrane fractions had been prepared. These data suggest that not all the α -adrenoreceptors present in rabbit uterus may be involved in mediating contraction induced by noradrenaline but do not define the nature of the additional α -receptors detected².

Using appropriate selective α -adrenoreceptor agonists and antagonists we have recently demonstrated in a functional test system that human platelets carry both α_1 - and α_2 -adrenoreceptors³. In appropriate conditions these cells can therefore be used to analyse the selectivity of α -agonists and antagonists for which this property has not been defined. In the case of an antagonist, selectivity is most appropriately analysed

by comparing the potency of the drug as an inhibitor of the clonidine (α_2 -adrenoreceptor-mediated) and methoxamine (α_1 -adrenoreceptor-mediated) stimulation of the response to ADP. When such studies are carried out using DHEC, the data obtained (Fig. 1) clearly demonstrate that this antagonist has very little ability to discriminate between the platelet α_1 - and α_2 -adrenoreceptors although it seems to be slightly more potent as an inhibitor of the α_1 -mediated response. The data of Fig. 1 closely resemble those obtained in similar studies using the classical non-selective α -adrenoreceptor antagonist, phen-tolamine. Peroutka *et al.*⁵ have reported that DHEC is also non-selective with respect to the proposed 'agonist' and 'antagonist' conformations of the α -adrenoreceptor in calf brain.

If the lack of α -selectivity of DHEC as defined here for human platelets is applicable to other tissues, as seems to be the case for uterine smooth muscle², caution should be exercised when this ligand is used to detect a correlation between a specific biological effect and receptor content. Thus, if the tissue contains significant concentrations of both α_1 - and α_2 -adrenoreceptors, estimation of receptor number from the extent of ^3H -DHEC binding will give misleading results unless selective α_1 - and α_2 -antagonists are used to indicate specificity. Similar considerations apply in studies which report

changes in the number of α -adrenoreceptors present resulting, for example, from prior exposure of the cells to catecholamines⁶.

MICHAEL C. SCRUTTON
JACQUELINE A. GRANT

Department of Biochemistry,
University of London King's College,
Strand, London WC2, UK

1. Williams, L. T. & Lefkowitz, R. J. *Science* **192**, 791-793 (1976).
2. Kunos, G., Hoffman, B., Kwok, Y. N., Kan, W. H. & Mucci, L. *Nature* **278**, 254-256 (1979).
3. Grant, J. A. & Scrutton, M. C. *Nature* **277**, 659-661 (1979).
4. Pearce, P. H., Wright, J. M., Fgan, C. M. & Scrutton, M. C. *Eur. J. Biochem.* **88**, 543-555 (1978).
5. Peroutka, S. J., Greenberg, D. A., U'Prichard, D. C. & Snyder, S. H. *Mol. Pharmacol.* **14**, 403-412 (1978).
6. Cooper, B., Handin, R. I., Young, L. H. & Alexander, R. W. *Nature* **274**, 703-706 (1978).

Are cardiac muscle cells skinned by EGTA or EDTA?

MILLER¹ re-examined criteria that have been used in defining the state of frog cardiac membranes treated with EGTA or EDTA. From these data, he concluded that 'supposedly' skinned cardiac cells are not skinned by EGTA/EDTA treatments.

To assess whether a preparation is skinned or not, it is helpful to have an objective definition of a skinned fibre preparation. This is especially necessary when using chemical methods of preparation that do not, as in mechanically skinned preparations, remove the major portion of sarcolemmal membranes. A definition we have found useful is that a fibre is skinned when normally impermeable solutes such as MgATP, EGTA, EDTA and other large molecular weight ions gain free access to the myofilament space.

Under this definition, critical physiological tests to determine whether a fibre is skinned must include one or more of these normally impermeable solutes, and a means for evaluating whether the solute has gained free access to the myofilament space. Examples that meet requirements of a 'critical test' for whether a preparation is skinned are (1) development of rigor and relaxation from rigor should occur within seconds of MgATP removal or addition, respectively (time interval predicted for ATP diffusion into or out of the fibre core)²; (2) the response of chemically skinned fibres to variations in low MgATP concentrations (0.1-50 μM) should be identical to that of mechanically skinned fibres and consistent with data³ obtained on extracted myofibrillar preparations; (3) the responses of chemically skinned fibres measured in (1) or (2) should not be quantitatively affected either by mechanical removal of the sarcolemma or by treatment of the pre-

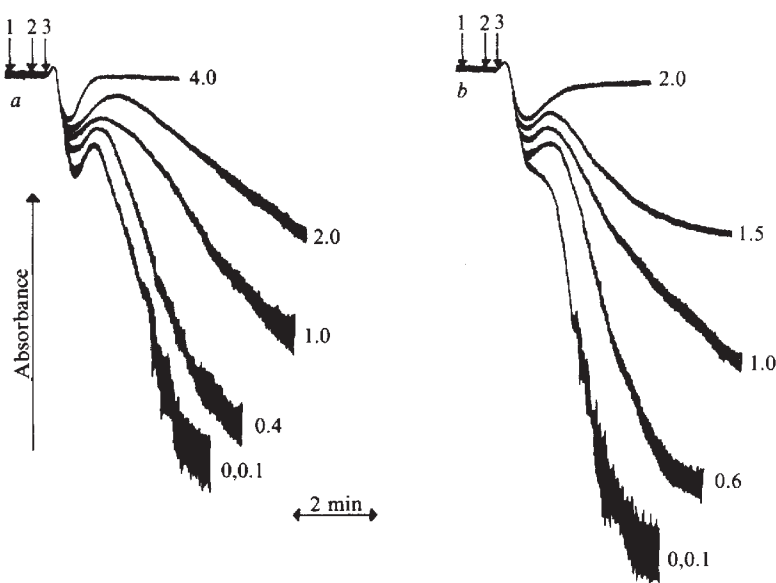


Fig. 1 Inhibition by DHEC of the stimulation of the response of human platelets to ADP induced by clonidine (a) or methoxamine (b). Aggregation was measured in 0.25-ml samples of human platelet-rich plasma prepared as described previously⁴. Additions were made as follows: at 1, DHEC at the concentrations (μM) indicated by the numbers on the trace; at 2, 5.7 μM clonidine (a) or 80 μM methoxamine (b); at 3, 1 μM ADP. Unpublished studies similar to those published previously³ demonstrate that methoxamine stimulation of the response to ADP is mediated by an α_1 -adrenoreceptor. DHEC at concentrations up to 5 μM had no effect on the response to 1 μM ADP added alone.

paration with non-ionic detergents that destroy the sarcolemma⁴.

We therefore suggest that a third criterion for determining whether a preparation is chemically skinned should be added to the two used by Miller of sensitivity of the muscle to Ca and lack of resting potential: that a preparation is chemically skinned if tension responses affected by MgATP are consistent with tension responses of mechanically skinned or non-ionic detergent-treated preparations. Because the behaviour of EGTA-treated mammalian skeletal muscle fibres⁴⁻⁷ meets this criterion, there can be no doubt that they come under the definition of a chemically skinned preparation⁵.

JOHN P. REUBEN

DONALD S. WOOD

Laboratory of Muscle Physiology,
Department of Neurology,
Columbia University,
New York, New York 10032

1. Miller, D. J. *Nature* **277**, 142 (1979).
2. Kushmerick, M. J. & Podolsky, R. J. *Science* **166**, 1297 (1969).
3. Weber, A., Herz, R. & Reiss, I. *Biochemistry* **8**, 2266 (1969).
4. Katz, G. M., Sorenson, M. M. & Reuben, J. P. *J. gen. Physiol.* **72**, 651 (1978).
5. Wood, D. S., Zollman, J., Reuben, J. P. & Brandt, P. W. *Science* **187**, 1075 (1975).
6. Wood, D. S., Sorenson, M. M., Eastwood, A. B., Charash, W. E. & Reuben, J. P. *Neurology* **28**, 447 (1978).
7. Reuben, J. P., Wood, D. S. & Eastwood, A. B. in *Pathogenesis of Human Muscular Dystrophies* (ed. Rowland, L. P.) 259 (Excerpta Medica, Amsterdam, 1977).

MILLER REPLIES—Reuben and Wood have defined a critical test of whether a muscle fibre may be considered skinned by treatment with EGTA or EDTA. This test provides an admirable basis against which to check all types of fibre, in addition to the skeletal fibres on which their own work has been carried out. However, their criteria differ markedly from those of Winegrad¹ which are still generally accepted (for review see ref. 2) and applied³, and which I contend to be ambiguous⁴. Indeed, additional criteria, such as the acceleration of the rate of tension development at higher Ca-buffer levels¹, are also met by unskinned preparations⁵ where extracellular exchange is the limiting process.

My own investigations on frog heart (on ventricular muscle, with which the technique was established¹, and auricle) demonstrated that (1) the method fails to produce the effects claimed for it, and (2) there are alternative explanations for the observed phenomena⁴. Subsequent work with the EGTA/EDTA method has, therefore, been founded on an incorrect interpretation and its validity ought to be re-examined in that light. Where additional evidence of skinning, preferably along the lines indicated by Reuben and Wood, has not been provided, the conclusions remain potentially open to criticism. This is not to say that all such preparations will prove, like frog heart, not to have been skinned, but that evidence broadly based on Winegrad's original criteria alone is not compelling.

Wood *et al.*⁶ had reported ultrastructural evidence of skinning by EGTA in skeletal muscle and the detailed results published more recently⁷ amply confirm their evidence of skinning. However, the feature of the method responsible for the skinning effect remains obscure as prolonged exposure to EGTA (80 mM) in itself does not lead to membrane disruption^{8,9}.

D. J. MILLER

1. Winegrad, S. *J. gen. Physiol.* **58**, 71 (1971).
2. Fabiato, A. & Fabiato, F. *Circulation Res.* **40**, 119 (1977).
3. McClellan, G. B. & Winegrad, S. *J. gen. Physiol.* **72**, 737 (1978).
4. Miller, D. J. *Nature* **277**, 142 (1979).
5. Miller, D. J. *Pflügers Arch. ges. Physiol.* **359**, 46 (1975).
6. Wood, D. S., Zollman, J., Reuben, J. P. & Brandt, P. W. *Science* **187**, 1075 (1975).
7. Katz, G. M., Sorenson, M. M. & Reuben, J. P. *J. gen. Physiol.* **72**, 651 (1978).
8. Barrett, J. N. & Barrett, E. F. *Science* **200**, 1270-1272 (1978).
9. Lüttgau, H. Ch. & Spiecker, W. *J. Physiol., Lond.* (in the press).

MILLER¹ has challenged the notion that the permeability of the sarcolemma of cardiac muscle can be markedly enhanced by treating the tissue with either EDTA² or EGTA^{3,4}. He has claimed that the two major criteria generally used to infer the existence of the high permeability state, that is, low transmembrane potential and increase in force at very low Ca concentrations, can be produced without an accompanying increase in membrane permeability, and he has concluded that only microdissection of the sarcolemma⁵ can produce a cardiac cell in which one can be sure that the membrane barrier to the movement of ions has been removed and the sarcoplasmic reticulum and mitochondria retained.

Although Miller's conclusions may apply to his own studies on EDTA-treated frog ventricle¹, it is less clear that they apply to my EDTA-treated frog ventricular preparation, as I found that the resting potential was not restored for many minutes after the re-introduction of normal Ringer's bathing solution (compare ref. 1, Fig. 1 with ref. 2, Table 1). In addition, in my preparation at pCa 5.4, tension was almost 90% of maximum whereas in Miller's it was less than 50% of maximum in spite of the same Na concentration. Miller's results do highlight the difficulty of producing a hyperpermeable state in frog heart with EDTA and focus on the reason for developing a better preparation.

Miller's conclusions are incorrect for hyperpermeable cardiac fibres that can be produced in rat ventricular muscle by treatment with EGTA as there are at least six different properties of this preparation that favour the existence of a high membrane permeability to small ions and molecules.

(1) The amount of force and the rate at which the force develops at any given concentration of Ca are very stable and reproducible⁴.

(2) A stable rigor state is rapidly initiated when ATP has been removed from

the bath and rapidly reversed when ATP is restored. This occurs in spite of the maintenance of the concentrations of Na and K and the elevation of the concentration of imidazole to keep the ionic strength constant. In the absence of ATP, the addition of phosphocreatine rapidly relaxes the rigor state only if ADP is present⁴. The relaxation therefore does not depend on a change in the concentration of Na.

(3) The membrane potential of individual cells was stable at -8 mV for long periods when all the KCl had been replaced with NaCl as well as in the presence of Ca concentrations as high as 0.03 mM. According to Miller and Mörchén⁶, reduction of Ca to 0.02 mM prolongs the action potential in frog only 20-200%, with little change in resting potential. In rat ventricle Garnier *et al.*⁸ found the duration of the action potential prolonged to only a few seconds by Ca-free solutions. Because in our study the same cell was impaled for several minutes, the depolarised state could not be due to a long plateau of an action potential.

(4) The Ca-induced Ca releases from the sarcoplasmic reticulum that were produced by lowering the concentration of EGTA to 30 μ M and raising Ca to sub-threshold concentration were very rapidly terminated by raising the EGTA concentration to 3 mM⁴. The very large increase in tension produced by caffeine-induced Ca release from the sarcoplasmic reticulum in 0.03 mM EGTA was totally blocked by 3 mM EGTA⁴.

(5) Although treatment of the EGTA-treated cells with detergent for 30 min destroys the function of the cell membranes, it does not alter either the rate of development of force at similar levels of tension or the maximum Ca-activated force.

(6) Possibly most important in refuting Miller's argument is the fact that the amount of force generated at a given concentration of Ca is not altered by changing the Na concentration in the bathing solution. The Na concentration was varied from 10 mM to 40 mM over a range of pCa of 9.0-4.5 with no change in the force generated at any concentration of Ca. These results indicate that Ca, ATP and EGTA all cross the membrane very rapidly and that gradients of K and Cl cannot be maintained across the membrane.

We have described the EGTA-treated cell as hyperpermeable because it is not the same as a mechanically skinned cell. The cell membrane remains, and it is still a diffusion barrier for large molecules such as creatine phosphokinase and glutamate oxaloacetate transaminase. This is not a limitation, but rather an advantage, as it allows the direct probing of the contractile proteins with Ca in a preparation that more closely resembles the intact cell, and regulatory mechanisms that are lost in the mechanically skinned fibre are retained in

the EGTA-treated hyperpermeable cells.

SAUL WINEGRAD

Department of Physiology,
School of Medicine,
University of Pennsylvania,
Philadelphia, Pennsylvania 19104

1. Miller, D. J. *Nature* **277**, 142–143 (1979).
2. Winegrad, S. J. *gen. Physiol.* **58**, 71–93 (1971).
3. McClellan, G. & Winegrad, S. *Nature* **268**, 261–263 (1977).
4. McClellan, G. & Winegrad, S. J. *gen. Physiol.* **72**, 737–764 (1978).
5. Fabiato, A. & Fabiato, F. J. *Physiol., Lond.* **249**, 469–495 (1975).
6. Miller, D. J. & Mörchen, A. J. *gen. Physiol.* **71**, 47–68 (1978).
7. Winegrad, S. J. *gen. Physiol.* **62**, 693–706 (1973).
8. Garnier, D., Rougier, O., Gargouil, Y. M. & Coraboeuf, E. *Pflügers Archs ges. Physiol.* **313**, 321–342 (1969).

MILLER REPLIES—To answer Winegrad's conclusion that frog heart is rendered 'hyperpermeable' by EDTA/EGTA treatment: the discrepancy in time scale for the restoration of normal resting potential between his study and my own^{2,3} can have several causes. First, my preparations are considerably smaller, are rapidly superfused and display tension responses within seconds^{2,4} rather than minutes¹ of solution changes. Slow extracellular exchange will contribute to delayed resting potential recovery. Second, Winegrad reported on the recovery of the resting potential after prolonged exposure to the full 'skinning' medium. It seems most likely that the intracellular ionic levels could alter in these circumstances (high $[K]_0$, low $[Na]_0$, depolarisation, low $[Ca]_0$) and contribute to the slow redevelopment of resting potential in normal Ringer's solution. Third, Winegrad's results were obtained with "an exploring microelectrode" (ref. 1, p. 79). This makes a reliable measurement of time course difficult as substantial repolarisation must occur to give a clear indication of a successful impalement with the microelectrode.

The differences in tension levels at pCa 5.4 noted by Winegrad are hardly indicative of qualitative differences between preparations. He has reported elsewhere that tension developed at this (or any other) pCa is strongly influenced by the immediate history of the preparation (for example, see ref. 5, Fig. 4). Quantitative differences among frog species⁴ and between auricle and ventricle (ref. 2 and my unpublished observation) do occur. However, in thin, rapidly superfused ventricle trabeculae exposed to a regimen identical to that of Winegrad, rapid ($t_{1/2} < 5$ s), complete relaxation always occurs when $[Na^+]_0$ is raised to 120 mM and $[K^+]_0$ reduced to 3 mM (pCa 5.4). This finding signifies more than the "highlighting of the difficulty of producing a hyperpermeable state".

I will discuss the six properties which Winegrad considers "favour the existence of a high membrane permeability to small

ions and molecules" in rat ventricular muscle. Properties (1) and (3) have been dealt with already (see above and refs 2, 3). With respect to property (2), removal of ATP as reported in Winegrad's publications had seemed to be associated with Na removal (ATP present as $Na_2ATP^{1,5-7}$). In experiments of the type that I previously reported², removal of Na_2ATP evokes a submaximal transient tension development (direct comparison is difficult as Winegrad's records are not calibrated in absolute or relative tension). After this the preparation becomes insensitive to pCa steps^{2,3}. The tension transient is blocked if NaCl (10 mM) is included in the ATP-free medium. These effects are also seen in frog trabeculae exposed to EGTA treatment overnight. Clearly, ATP removal does not produce rigor tension in frog. These comments are not necessarily appropriate to Winegrad's experiments in the rat. However, two points can be made.

First, tension development in these conditions is not necessarily rigor tension. Stiffness measurements would form one potentially definitive line of evidence for or against the occurrence of rigor and hence the hyperpermeable state. Second, if the rat cells are not 'hyperpermeable' then the effects of ATP removal, and ADP and phosphocreatine addition⁶, would be extracellular. Extracellular ATP was responsible for qualitative changes in the behaviour of the frog preparations². The concentrations of ATP ($MgATP^{2-}$, ATP^{4-} ?) involved await detailed analysis. ADP and phosphocreatine have yet to be investigated, but they could act in tandem outside the cell. However, as Winegrad has noted⁷, leakage of creatine kinase is increasingly used as an index of myocardial damage. This raises the possibility that a low, but significant, level of extracellular ATP is produced by both nonspecific sites and by creatine kinase leakage from damaged cells. Certainly, this type of experiment is potentially capable of differentiating between skinned and non-skinned cells (see also Reuben and Wood's comments above).

To discuss property (4), even with rapid superfusion of very small preparations, ≥ 0.1 mM EGTA (in nominally Ca-free media) is required to prevent tension development in Na-poor solutions⁴. Tension oscillations in these poorly Ca-buffered conditions cannot readily distinguish between Ca-induced Ca-release phenomena in intact and skinned cells as they are triggered by local $[Ca^{2+}]$ fluctuations in both cases: extracellular $[Ca^{2+}]$ near the sarcolemma for intact cells and $[Ca^{2+}]$ adjacent to the sarcoplasmic reticulum for skinned cells. Extracellular $[Ca^{2+}]$ steps produce Ca release in K-depolarised frog heart⁸ and oscillatory responses have been reported⁹.

Referring to property (5), the comparability of both the rate of tension development and the maximum force in

EGTA- and detergent-treated cells is not a definitive property. Ca-exchange would be limited by the gross morphology of the preparations whether the cell membranes are 'hyperpermeable' or not (see also first paragraph). However, the Ca-sensitivity of detergent-treated preparations differs markedly from that of EGTA-treated cells, half-maximum tension occurring more than 1 pCa unit lower in the latter^{6,7}. Winegrad attributes this to a modifying effect of fragments of the sarcolemma persisting after EGTA treatment and subsequently removed by detergents. This idea contrasts markedly with the criteria suggested by Reuben and Wood (see above). Similarly, Winegrad's interpretation of the effects of 5 mM theophylline as occurring through phosphodiesterase inhibition⁷ rather than Ca release from sarcoplasmic reticulum (like caffeine and other methylxanthines^{10,11}) seems unduly selective.

Considering property (6), the sharp sensitivity of EDTA/EGTA-treated frog heart to $[Na^+]$ provides a clear-cut distinction from that of skinned fibres. In frog, the curve relating tension to $[Na^+]$ is shifted more than one order of magnitude towards higher $[Na^+]_0$ by Winegrad's media compared with 'simple' Na-poor solutions (see ref. 2, Fig. 3 and ref. 4, Fig. 4). As these effects have not been quantified for the mammal one cannot assume that any particular $[Na^+]$ will produce relaxation, but levels higher than 40 mM are worth trying. (Note: $K_{Na}^{app} \cdot [Na]_0^4$ in Fig. 3, ref. 3, is given as

$$\frac{1}{K_{Ca}[Ca]^2} + \frac{K_{Na}}{K_{Ca}[Ca]^2} \cdot [Na]^4$$

which does not significantly affect the shape of the curve.)

The failure of the EGTA 'skinning' technique has been demonstrated for frog heart, and a plausible mechanism presented to account for the findings which had proven misleading². The criteria for successful skinning as previously applied to frog are, therefore, inadequate. As a result, a definition of skinning in other tissues is not at all straightforward, but the ideas presented by Reuben and Wood should form a useful starting point.

D. J. MILLER

Institute of Physiology,
University of Glasgow, UK

1. Winegrad, S. J. *gen. Physiol.* **58**, 71–93 (1971).
2. Miller, D. J. *Nature* **277**, 142–143 (1979).
3. Miller, D. J. & Mörchen, A. J. *gen. Physiol.* **71**, 47–67 (1978).
4. Miller, D. J. & Moisescu, D. G. J. *Physiol., Lond.* **259**, 283–308 (1976).
5. Winegrad, S. J. *gen. Physiol.* **62**, 693–706 (1973).
6. McClellan, G. & Winegrad, S. *Nature* **268**, 261–263 (1977).
7. McClellan, G. & Winegrad, S. J. *gen. Physiol.* **72**, 737–764 (1978).
8. Miller, D. J. J. *Physiol., Lond.* **242**, 93–95P (1974).
9. Winegrad, S. J. *gen. Physiol.* **68**, 145–157 (1976).
10. Chapman, R. A. & Miller, D. J. J. *Physiol., Lond.* **242**, 615–634 (1974).
11. Miller, D. J. & Thieleczek, R. J. *Physiol., Lond.* **273**, 67–68P (1977).

reviews

General Relativity after Einstein

Robert Wald

General Relativity: An Einstein Centenary Survey. Edited by S. W. Hawking and W. Israel. Pp. 831. (Cambridge University Press: Cambridge, 1979.) £37.50.

MARCH 14, 1979, marked the one hundredth anniversary of the birth of Albert Einstein and numerous activities commemorating this event have taken place or are planned for the near future: the erection of a statue, the convening of conferences and the publication of memorial volumes, such as this book. I do not wish to speculate on how Einstein would have viewed all these activities; in any case, in the final analysis their value must be judged on their own scientific (or other) merits, not on their proclaimed connection to Einstein. Judged in this light, the Hawking-Israel volume is, in all respects, a marvellous book.

When two distinguished scientists edit a book consisting of sixteen articles contributed by other distinguished scientists, it is almost impossible for a bad book to be produced. But the Hawking-Israel volume is quite a bit more than merely "not a bad book". The articles cover the complete spectrum of progress in relativity theory in the past decade, and almost all are very well written. Most of them should be accessible to graduate students in relativity and yet almost all contain some new results or new points of view that should be of interest to the specialised researcher. Of course, no book is perfect and in this volume there is some disparity from article to article in the technical expertise expected of the reader. But this fault — which is almost unavoidable in a book of this nature — is about the only one important enough to be worthy of mention.

Very little work on general relativity was done during the 1940s and 1950s, but there has been a great deal of activity and progress since then, both in development of the theory and in experimental (or observational) verification of its predictions. With regard to the latter, for many years the empirical basis of general relativity rested on the fact that its predictions were in accord with observations in the three classical tests: the gravitational redshift, the bending of light, and the precession of Mercury's orbit. In recent years, however, the use of techniques such as long baseline radio interferometry (to measure the gravitational deflection by the Sun of radio signals from a quasar) and radar (and laser) ranging (to measure gravitational time delay of signals and

determine orbits of bodies) have greatly improved measurement accuracy; and the discovery of a pulsar in a close binary orbit has given us an apparently clean system where effects of strong gravity are important and can be studied. In the meantime, a theoretical frame has been developed for analysing alternative theories of gravitation and determining which aspects of each theory the various experiments and observations actually test. All these developments are comprehensively reviewed in the chapter by Will, who concludes that, so far, general relativity has passed all empirical tests, but more stringent tests will confront it in the near future.

One important prediction of general relativity that has not yet been directly confirmed (although it can now be inferred from changes in the orbit of the binary pulsar) is the existence of gravitational radiation. The chapter by Douglass and Braginsky estimates the characteristics of gravitational radiation expected from the various possible astrophysical sources and discusses the schemes for detecting this radiation. The problems of detection are formidable, as even the radiation from a supernova within our galaxy is estimated to produce fractional relative displacements of only $\sim 10^{-17}$. Nevertheless, the authors are optimistic that an unambiguous discovery of gravitational radiation will occur in the near future.

A key ingredient in the renaissance of theoretical work in general relativity beginning in the 1960s was the use of techniques of modern differential geometry. These techniques were used to prove theorems about the global structure of spacetime, in particular the inevitability of singularities in certain conditions and important results in the theory of black holes. These 'global methods' are reviewed in the chapter by Geroch and Horowitz. Their review focuses on the basic ideas and issues and illustrates how a theorist goes about proving a theorem or constructing a counterexample. The 'expert' who wants complete proofs of the strongest known theorems would be best off consulting other sources, such as the textbook of Hawking and Ellis. But the non-expert — or the graduate student who wishes to become an expert — should find this article to be the best available introduction to the subject.

Progress in the analysis of Einstein's equation is the subject of the chapter by Fischer and Marsden. The main topics

of their review are the initial value formulation of general relativity, linearisation stability (that is, the question of whether a linear perturbation corresponds to an exact solution), and other issues involving the manifold structure of the space of solutions. In contrast to the Geroch-Horowitz chapter, this chapter is written for the expert and is rather demanding of the reader. Most relativity theorists will also have difficulties following the notation, as index notation for tensors is not used.

One of the most exciting areas of progress in the past decade has been the full development of the theory of the black hole, the "region of no escape" which results from the gravitational collapse of a body. Among the marvellous results that have been obtained are the proof of uniqueness of black hole stationary states, the theory of energy extraction, and the analogy between the laws of black hole mechanics and the laws of thermodynamics. These and other results are reviewed in the chapter by Carter.

The marvellous properties of black holes do not end with the above general theory. Studies of the equations governing gravitational, electromagnetic, and other perturbations of the Kerr solution — the unique stationary black hole — have shown that decoupled equations can be obtained and can be solved by separation of variables. Furthermore, numerous differential identities exist between solutions. The chapter by Chandrasekhar provides a thorough introduction and summary of his work on this problem.

The role of black holes in astrophysics is the subject of a very well written review by Blandford and Thorne. These authors summarise the theoretical scenarios where black holes may occur and the astrophysical effects they would produce, as well as the observational evidence for them. In addition, Blandford and Thorne do a good job in conveying a sense of the nature of the assumptions which go into the various models. Thus, the relativity theorist or other non-astrophysicist who wants to know how seriously to take reports of discoveries of black holes will find this chapter to be a valuable source.

The other area besides black holes where the effects of general relativity play a dominant role is cosmology. A brief chapter by Dicke and Peebles presents some interesting — though rather

sketchy — speculations on issues related to the apparently finite origin of our Universe. Zel'dovich describes the history of the early Universe, including discussions of primordial black holes, quantum effects, and implications of particle physics for cosmology (and *vice versa*). The chapter by MacCallum provides a thorough review of homogeneous, anisotropic cosmological models as well as the rather meagre known results concerning inhomogeneous cosmology.

One of the highlights of the book is Penrose's beautifully written, thought provoking chapter on time asymmetry. Penrose considers the much discussed question of the origin of the arrow of time, but his ideas are quite original, though, as he continually points out, well within the realm of conventional physics. Penrose's main thesis is that at least some, and possibly all, of the various arrows of time — such as entropy increase and the retardation of radiation — can be explained by assuming the initial state of the Universe is one of low entropy. He speculates that this conclusion should be incorporated into physical law, perhaps in the form: The Weyl tensor vanishes at all initial singularities. (The Weyl tensor is a rough measure of gravitational clumping, so the intent of the proposed law is to state that initial gravitational entropy is small.) Whether or not this speculation is correct, the reader should find Penrose's discussion of cosmic censorship, his criticisms of Hawking's "black hole is a white hole" hypothesis and his fundamental criticism of chaotic cosmology, as well as numerous other side discussions, to be of great interest.

The final four chapters of the book deal with quantum aspects of general relativity. Virtually all researchers agree that classical general relativity cannot be an exact description of nature; that the gravitational field, like all other fields, must be quantised. But major conceptual issues arise when one attempts to do so, resulting mainly from the fact that the quantity which describes the gravitational field — the spacetime metric — also describes the background structure of spacetime. If one ignores this issue and tries to quantise gravity by procedures analogous to those used for other fields, one finds that the theory is non-renormalisable, that is, in perturbation theory new terms must be introduced in order to cancel the ultraviolet divergences. Presumably, an infinity of new, undetermined, parameters would appear if one went to arbitrarily high order in perturbation theory.

These difficulties have been known for a long time, but two recent developments have made theorists considerably more optimistic that they can be overcome. The first is the success achieved in particle physics in quantising gauge theories and unifying the fundamental interactions. Gravity would fit very naturally into

further developments of this programme. The second is Hawking's calculation of thermal particle creation by black holes due to quantum effects. The beauty of this result has provided the psychological equivalent of experimental verification that the theory is on the right track.

Almost all of the solidly grounded results on quantum effects in strong gravitational fields have been obtained in the approximation where the metric can still be treated classically. These results include calculations of particle creation (such as Hawking's black hole calculation) and 'back-reaction' of the particles on the gravitational field. This work is nicely reviewed from the Cambridge viewpoint by Gibbons. The chapter by DeWitt also contains an extensive review of this work, including a good discussion of the observer dependence of particle detection. Proposals for dealing with gauge problems in full quantum gravity are also discussed by DeWitt.

The chapter by Hawking is probably the best available introduction to his current ideas on quantum gravity. Hawking uses the path integral approach and proposes a solution to the problem of the indefiniteness of the gravitational action (which causes convergence problems in the path integral). His ideas for going beyond the lowest approximation lead him to a picture of "spacetime foam", a highly complex topological structure of spacetime on Planck length (10^{-33} cm)

scales. Much of this work is highly speculative and the logic is not always airtight; but the beauty of many of the ideas should convince the reader that there must be much fundamental truth contained in this work.

The chapter by Weinberg takes a much more conventional viewpoint on quantum gravity. Weinberg reviews the evidence that quantum gravity is non-renormalisable, but points out that if the condition of "asymptotic safety" (defined and discussed in the chapter) holds for the theory, then all but a finite number of the infinity of parameters occurring because of non-renormalisability could be fixed in a natural way. This would tame the non-renormalisability problem and could make the theory acceptable without further modifications. Some evidence that quantum gravity might be asymptotically safe is provided by a model calculation in $2 + \epsilon$ dimensions.

As the above discussion should indicate, the Hawking-Israel volume provides a dramatic demonstration of recent progress and future prospects of general relativity theory. The time is ripe for a book of this kind, and it is fortunate that the anniversary of Einstein's birth has provided the impetus for publishing it.

Robert Wald is Assistant Professor in the Department of Physics and in the Enrico Fermi Institute at the University of Chicago, Illinois.

Control of prokaryotic gene expression

The Operon. Edited by J. H. Miller and W. S. Reznikoff. Pp. 449. (Cold Spring Harbor Laboratory: Cold Spring Harbor, New York, 1978.) \$42.

The Operon is a collection of review articles that summarise much of what is known about the control of prokaryotic gene expression. The first half of the book describes aspects of regulation in the *Escherichia coli* lactose operon: the system that is best understood. The remainder of the book focuses on other systems which contrast with the *lac* operon: regulation by temperate phage λ repressor, the regulation of λ repressor synthesis, dual control of the galactose operon, the synthesis of 3',5' cyclic AMP, positive regulation of the arabinose operon, attenuation control of the tryptophan biosynthetic operon, autogenous regulation of the histidine degradation operon, and the mechanism of phase variation in *Salmonella*. The book is highly successful in that the articles are well written and provide a very broad education. It would be suitable for graduate courses and useful as a source of

information for instructors in molecular genetics.

Jonathan Beckwith's first article describes the genetic methods used to study the *lac* operon. This chapter is sufficiently clear and general to allow the reader to understand microbial genetics in a broader sense. Beckwith provides a second chapter on methods for fusing the β -galactosidase gene into foreign regulatory units, so that the simplicity of the β -galactosidase assay can be used to probe very complex regulatory systems as well as the mechanism by which proteins are exported from the cytoplasm to the cell membrane and into the medium.

Jeffrey Miller's chapter gives a concise history of genetic studies on the lactose repressor (*lacI*) gene. It also contains a detailed description of the genetic methods used to create mutant repressors with interesting physical and chemical properties. An extra attraction of this article is its summary of how the mechanism of mutagenesis was worked out, using the known DNA sequence of the *lacI* gene, in combination with a variety of genetic manipulations.

The chemical properties of normal and mutant *lac* proteins are thoroughly discussed by I. Zabin and K. Beyreuther. I was interested to learn that both the specific and nonspecific DNA-binding properties of the *lac* repressor are

confined to the amino-terminal 53 amino acid residues, whereas the ability to form tetramers and to bind inducer requires the carboxy-terminal part of the repressor molecule. In a similar fashion, λ phage repressor seems also to contain two domains: an amino-terminal DNA-binding region and a carboxy-terminal domain needed for oligomerization and derepression.

Terry Platt has contributed a fine article on the tryptophan biosynthetic operon, which is regulated by a combination of repression at the operator and antitermination at the nearby 'attenuator'. Transcription termination provides a versatile and sensitive form of regulation, as the *trp* attenuator responds

to changes in nucleotide sequence, to termination factor *rho* and even to the structure of tryptophan tRNA. Platt mentions the fact that the histidine biosynthetic 'operon' has no repressor nor operator and is controlled almost entirely by antitermination. Thus, the histidine operon would be worthy of a separate article.

The chapter on phase variation provides the greatest novelty in the book, as the inversion of several hundred nucleotide base-pairs determines whether the H-1 or H-2 flagellar antigen is synthesised.

Richard Calendar

Richard Calendar is Professor of Molecular Biology at the University of California at Berkeley.

Sexuality in man and animals

Sex, Hormones and Behaviour. Ciba Foundation Symposium No. 62. (New Series.) (Excerpta Medica: Amsterdam, Oxford and New York, 1979.) \$41; Dfl.84.

THE purpose of this symposium was to explore, and extend, the zone of overlap between studies of human sexuality and experimental research on sexual behaviour in animals. The book contains fourteen papers, dealing mainly with aspects of sexual development, homosexuality, and the neuroendocrine regulation of heterosexual behaviour.

In many respects, the overlap between human and animal research in these areas is extensive and this is particularly

apparent where studies of the non-human primates are concerned. The degree to which androgens and experiential factors influence the development of sex differences in human behaviour remains a controversial issue, as indicated by contributions from Ehrhardt, Meyer-Bahlburg and Green. Fresh insights into these problems could emerge from research on monkeys. Abbott and Hearn describe experiments on common marmosets which indicate that sexual differentiation of the brain by androgen occurs, at least partly, during postnatal life. In future it may be possible to use marmosets as models to study how androgens affect the developing brain of male primates.

Very little is known about the neuroendocrine control of human sexual behaviour in adulthood, or the extent to which steroid hormones, brain monoamines and peptides influence such

behaviour. Contributions by Nieschlag, Bancroft and Skakkebaek describe the effects of androgens on human sexual behaviour, and Everitt presents new findings concerning monoamines and sexual behaviour in rhesus monkeys. The interplay between social factors, neuroendocrine mechanisms and sexual behaviour constitutes an important field of research which is explored by Keverne, in his paper on laboratory groups of talapoin monkeys.

Satisfactory animal models with which to study certain facets of human sexuality have yet to be defined. Homosexuality and transsexuality are specific examples. It is evident that approaches to the study and treatment of homosexuality advocated by Dörner at this symposium were not endorsed by some participants. In his paper, Beach provides thorough guidelines on the potential value and limitations of animals as subjects for research on homosexuality and heterosexual behaviour in man.

This is an excellent book. One minor criticism is that some contributors present purely qualitative papers, whereas the reader might find a quantitative treatment more useful. Discussion sessions between the participants are reported fully, as in other Ciba Foundation Symposia, and total almost 150 pages. These discussions greatly enhance and amplify the value of the papers, making this volume a valuable addition to the literature on sexual behaviour in man and animals.

Alan F. Dixon

Alan F. Dixon is a Research Fellow at the Wellcome Laboratories of Comparative Physiology, Zoological Society of London, UK.

Atomic absorption spectroscopy

Atomic Absorption Spectroscopy. Second edition. By M. Salavin. Pp.193. (Wiley: New York and Chichester, UK, 1979.) £14; \$26.

THIS monograph is the second edition of Volume 25 of a series on analytical chemistry and its applications edited by P.J. Elving, J.D. Winefordner and I.M. Kolthoff. As the first edition was written by Walter Slavin, son of the present author, Morris, they join that select company of father-and-son books so prized by collectors, but in reverse order.

In his preface the author states that this should be regarded as a new book on atomic absorption spectroscopy from a different perspective. The organisation and layout of the book, however, follow lines similar to the first edition but, naturally, in view of the developments that have occurred in the intervening eleven years, the material content of the sections has been brought up to date and in many

cases completely rewritten. The additional areas covered include a section on electrothermal atomisation and greater emphasis is given to the chemical preparation of samples and to environmental applications. The author rightly draws attention to the neglected area of high precision analysis for major elemental constituents but offers little fresh guidance on how to achieve it. An appendix on commercial instrumentation contents itself with a list of eleven sources of relevant commercial equipment of which only four are actual suppliers of atomic absorption instruments. The information in this section is cursory and refrains from mentioning any European manufacturers.

The different perspective follows inevitably from the author's obvious experience in optical emission spectroscopy which by the same token demonstrates his unfamiliarity with the conventional terminology of atomic absorption spectroscopy. This makes the book slightly upsetting for practitioners of the art but more importantly sets a poor example for the newcomer or for the bench worker in atomic absorption spectroscopy,

for whom the book is intended.

The book is written in an off-hand, chatty style, which would be refreshing enough were it not casual to the point of inaccuracy. The use of jargon from other disciplines — who outside the field of emission spectrometry knows what a direct-reader is? (p14) — is less than helpful. Many of the terms used are positively misleading, for example, the sample is "injected" (p11) and the flame is described as being "scanned" by a beam from a hollow-cathode discharge lamp (p9 and p27). Other errors are to be found — nitrous oxide is hardly a fuel (p11), reference 33 should be to Zaidel and in Table 3:1 (p53) 50% O₂-50% N₂ would hardly make a satisfactory flame.

The book is easy to read and can be regarded more as an essay on the subject than as a satisfactory laboratory handbook especially in a field where a plethora of excellent textbooks already exists.

Allan M. Ure

Allan M. Ure is a Principal Scientific Officer in the Department of Spectrochemistry at the Macaulay Institute for Soil Research, Aberdeen, UK.

Neural modelling

Neural Modelling: Electrical Signal Processing in the Nervous System. By R. J. MacGregor and E. R. Lewis. Pp. 474. (Plenum: New York, 1979.) \$30.

NEURAL MODELLING is an activity carried out by theoretical biologists, mathematicians, bioengineers, computer types and some physiologists, the aim of which is to provide insights into how brains work. As such, it has aroused considerable misgivings, to say the least, in the minds of many, if not most, experimental neurobiologists. One distinguished neurophysiologist maintains, perhaps with tongue in cheek, that neurobiology is "totally non-mathematical"; another that mathematical neurobiology is mostly wrong, long on formal theory, but short on facts, and that what is not wrong is trivial. One can perhaps understand, if not sympathise with such feelings, given that most neurobiologists are profoundly ignorant of advanced mathematics, and have formed a rather stereotyped view of the neurotheorist, perhaps not unjustified, as one not really interested in, nor capable of distinguishing 'real' facts.

It is against such a background that this book should be seen. MacGregor and Lewis have tried to cover everything from the biophysical properties of the neuronal membrane to the computational aspects of neocortical nets, in some fifteen chapters comprising 380 pages. Chapters 1 and 2 provide a quick introduction to the neurone doctrine, that the brain is made up of a coupled nerve cells interacting across specialised synapses, and to the notion of the unit compartment: that the critical element in modelling electrical signal processing in cells and in nets is a patch of membrane, rather than an entire cell.

Chapters 3 and 4 cover in some detail the biophysics associated with the generation of transmembrane resting potentials. Readers will find in here a good discussion of the Nernst-Planck and Goldman equations, of Donnan equilibria, of ionic and electrogenic pumps, and the like.

Chapters 5–7 follow this up through the Eccles equivalent circuit for postsynaptic membrane potential changes, the Fuortes–Hodgkin–Baylor models for photoreceptor transduction, Rall's work on cable electrotonus, the Hill–Monnier–Rashevsky theories of suprathreshold excitation, to the Hodgkin–Huxley equations, but not beyond.

There follows a radical change of pace and style, as the authors begin the treatment of neuronal nets. Chapter 8 describes current work on analogue

computer simulations of small nets of idealised neurons, called neuromines, and their utilisation in the modelling of invertebrate neuronal structures. Chapters 9 and 10 describe applications of stochastic process theory to model neuronal firing patterns, and the various statistical techniques which have been used to discern regularities in the firing patterns of single and paired cells. Chapters 11–13 deal with the dynamical properties of large-scale nervous activity, seen through Monte Carlo simulations, continuum field theories, or digital computer simulations. Chapter 14, which is very short, discusses briefly, slow potentials and the EEG.

Chapter 15 deals with a number of attempts to understand the functioning of such neuronal structures as the visual neocortex by Pitts and McCulloch, and by Marr; the generation of thalamic and hippocampal rhythms, by Anderson *et al.*; and the brain-stem reticular formation; and there is a brief mention of lateral inhibition in the retina. Finally, chapter 16 closes with a few methodological comments about modelling in general, and brain modelling in particular.

In my opinion, the authors have failed to provide a sufficiently penetrating account of the field of neuronal

modelling, up to mid-1977. It is, of course, always a problem in seeking to convey the nature of a discipline, to choose between breadth or depth. To my taste, only chapters 3 and 4 cover the material in sufficient depth as to be interesting. The others are so superficial as to be relatively boring for anyone who knows anything about the subject matter. In addition to this, I feel that the book is written from the wrong perspective, from the bottom up as it were, instead of from the top down. Over the past 5 years or so, it has become apparent to many neurotheorists, that trying to deduce what is going on in neuronal circuits from the properties of neurones and synapses is not very profitable, and that one has to understand in some detail the computational problems which have to be solved by the system and the algorithms used, before one can get to questions of how such algorithms are embodied in neuronal circuits. This is only hinted at in the later chapters of this book. A really good book on the subject of neuronal modelling remains to be written.

Jack Cowan

Jack Cowan is Professor of Theoretical Biology at the University of Chicago, Illinois.

Chromatographic methods

Laboratory Handbook of Chromatographic and Allied Methods. Edited by O. Mikeš. Pp. 764. (Ellis Horwood/Wiley: Chichester, UK, 1979.) £38.50.

THIS is an excellent addition to the numerous practically oriented books in the Ellis Horwood Series in Analytical Chemistry. Dr Mikeš has enlisted the help of 13 of his colleagues, all employed in the Czechoslovak Academy of Sciences, to produce a completely rewritten version of his *Laboratory Handbook of Chromatographic Methods* (1966). I consider the present book to be a worthy successor to E. and M. Lederer's *Chromatography* (1953, 1957).

The "allied methods" alluded to in the title are countercurrent distribution and electromigration, and one chapter is allocated to each. The chapter on countercurrent distribution is clear about matters of principle. However, it ignores the dramatic usefulness of the technique in determining the structure of the antibiotic peptides and in the purification of transfer RNAs by Holley and colleagues. Moreover, no mention is made of Alderweireldt's approach, which culminated in the development of 'steady-state' machines; the example given of use of such a machine is undocumented and fails completely to illustrate its capabilities

for finding needles in haystacks. The chapter covering electrophoresis is unclear on matters of principle and inadequate in coverage.

By contrast, the 10 chapters relating to chromatography proper provide excellent coverage both of principle and of manipulative practice. Analytical applications described in the text are by way of example only, but the references to reviews and original papers make it easy to find out how any particular technique has been used to deal with particular classes of organic or inorganic substances. There are some notable omissions. Liquid-liquid chromatography on columns and preliminary concentration from crude mixtures by displacement are scarcely mentioned. Although carrier displacement is discussed in principle, no mention is made of Sanger's brilliant application of it to oligonucleotides. Indeed, there is little practical advice in the whole work about detecting or measuring radioactive substances in chromatography. However, in general, the treatment of chromatography is excellent and the practical detail and careful and extensive bibliography will make this book a good investment for the library of every establishment where chemical analyses are carried out, and particularly for those establishments where versatility is demanded.

R.L.M. Synge

R.L.M. Synge is Honorary Professor of Biology at the University of East Anglia, Norwich, UK.

obituary

Giulio Natta, 1903—1979

GIULIO NATTA, the eminent Italian Chemist, who died on 1 May 1979, at the age of 76, will be remembered for his discovery of isotactic polypropylene, and for his outstanding researches on the synthesis and structure of stereo-regular polymers. Seldom have such scientific contributions aroused such profound fundamental interest, and his work sparked a volume of research on polymerisation never before approached.

Natta was born in 1903 in Imperia, a small seaside town in Liguria. His father was a well known judge, but contrary to the traditional profession of his family, he did not take up law studies. He was fascinated by chemistry from his school days, and after attending the 'Christopher Columbus' school in Genoa, he studied mathematics at the University of Genoa. He changed disciplines when he moved to the Milan Polytechnic Institute where he took his 'Dottore degree' in chemical engineering at the early age of 21, and his 'Libero Docente' three years later. He was assistant lecturer in chemistry at Milan before he moved in 1933 to the University of Pavia, as Professor of General Chemistry, and two years later to Rome as Director of the Institute of Physical Chemistry. In 1937 he became Professor of Industrial Chemistry at Turin, and finally from 1938 until he retired in 1973, Professor and Director of the Milan Institute of Industrial Chemistry.

Natta's varied professional duties and wide range of research interests equipped him for the remarkable contributions he was to make to polymer science from 1954 onwards. He ascribed the particular conditions, which enabled his research school at Milan to achieve such rapid and conclusive results in the genesis and structure of new classes of macromolecules, to his early experience with X-ray studies of the structure of crystals and the resolution of chemical and structural problems.

At first his investigations were directed to the study of the structure of low molecular weight inorganic substances, but in 1932, following a chance meeting with Professor Hermann Staudinger, he was attracted to the study of the lattice structure of linear polymers. This aspect of his work did not develop very rapidly and he concentrated his X-ray work on the structure of heterogeneous catalysts

used for certain important industrial applications. At the same time he investigated the processes for the synthesis of methanol, higher alcohols and formaldehyde and in 1938 he was invited by the Italian Government and some industrial companies to initiate a research programme on the production of synthetic rubber. The industrial production of butadiene-styrene copolymer rubbers was realised in Italy within a few years and he also developed a purely physical commercial process for the separation of butadiene from butene-1. At the same time he began a series of researches on the applications of petroleum derivatives, and in particular the use of olefins and diolefins as raw materials for chemical synthesis such as the oxo-synthesis and polymerisation.

In 1952 he attended a lecture by Karl Ziegler¹, the Director of the Max Planck Institute at Mulheim, and was struck by the fact that on reacting an olefin with an organo-metallic compound, Ziegler obtained a dimerisation which resulted substantially in a single product and not a mixture. Afterward he said of the meeting "the knowledge acquired in the field of polymerisation enabled me to appreciate the singularity of the methods that Ziegler described and my interest was aroused".

Natta suggested to Montecatini, Italy's largest chemical company, for whom he was a consultant, that a direct link should be established with Ziegler, so that his work and ideas could be developed. The outcome was that Montecatini purchased rights for the commercial development of Ziegler's work in Italy and Natta acquired access to Ziegler's researches on the transformation of olefins and started researches in Milan on the polyaddition of ethylene with organic aluminium compounds.

In 1953 Ziegler made the momentous discovery of the polymerisation of ethylene to a high molecular weight linear polymer and immediately thereafter concentrated his research on the development of better and more efficient catalysts, whereas Natta, on the other hand was firmly convinced that the Ziegler catalyst contained the key to the regular polymerisation of olefins, dienes and other unsaturated compounds.

It was in March 1954 that he made his great discovery of the stereo-regular

polymerisation of propylene, butene-1 and styrene. Feverish research resulted within a short time in the polymerisation of many olefins to isotactic polymers, propylene to the syndiotactic polymer, ethylene-propylene copolymers and the stereo-specific polymerisation of non-hydrocarbon monomers. The work in his laboratory was organised so as to obtain maximum collaboration between the different disciplines, engineers worked alongside chemists and physicists in a group consisting of as many as 60 scientists of different levels and status.

Natta was the architect and dealt with all things which were to be done, or had to be, or could be done. In scientific matters he favoured experiment rather than theory, facts rather than hypotheses and relied very much on his exceptional intuition. In all these matters he combined the dedication of the pure scientist with the practical logic of the engineer.

About 1956, Natta developed the symptoms of Parkinson's disease. He accepted his illness as a fact of life and tried to continue as though it was not there. His wife and colleagues were a great help to him at this time, and when he received his Nobel Prize from the King of Sweden in 1963, he was supported by his son Guiseppe and his Nobel Lecture was read by a colleague. He gave up work finally in 1973 and lived near his daughter Franca in complete privacy.

Away from his laboratory he was a keen nature lover, and an active mountain climber. His main hobby was collecting fossils, which he successfully displayed in his home in Milan. Another hobby which he pursued for many years was collecting wild mushrooms and he had an extensive knowledge of edible fungi.

Those who knew Natta well attest to his brilliance of mind and exceptional capacity for work. He was a quietly dressed, soft spoken man of short stature and showed courtesy and consideration to visitors from all over the world who came to Milan, either out of scientific interest, or out of interest in the practical and industrial development of his discoveries. His work was internationally recognised by awards in many countries, and he was co-winner with Karl Ziegler of the 1963 Nobel Prize for Chemistry for his work on olefin polymers.

C.E.H. Bawn

1. *Obit. Not. Fell. R. Soc. Lond.* **21**, 569-584 (1975).

Elias Melin

WITH THE death of Elias Melin on 22 March 1979, at the age of 89, experimental botany and, more especially, mycorrhizal research has lost a pioneer who held a leading position within his field. His experimental life spanned a period of over 60 years, during which he kept abreast of new developments in his area. Indeed, up to a year before his death, he was still working slowly but productively in his laboratory. His interests were wide and covered aspects of ecology and physiology of higher plants and of fungi.

Melin started his scientific career as a student of the well known plant ecologist Rutger Sernander who in the early decades of this century attracted a great number of gifted students, creating a school of plant sociology which played an important role in the development of Scandinavian botany. At first Melin became interested in the taxonomy and biology of the genus *Sphagnum*, and his doctoral thesis (1917) comprised a study on the plant sociology of North Scandinavian peat bogs. During these studies he observed that young seedlings of pine and spruce grew well on drained peat bogs only if they became infected by mycorrhizal fungi. After this discovery, Melin devoted his life to studies on the mycorrhiza (the symbiotic association between tree roots and fungi) of forest trees.

As a university lecturer at the Royal College of Forestry in Stockholm, he described the structure and the fungal components of the mycorrhizas of pine, spruce, larch, birch, and aspen. He worked out methods for the isolation in pure culture of the fungal symbionts and for the synthesis of mycorrhiza of forest tree seedlings under controlled laboratory conditions. He proved that various species of *Boletus*, *Amanita*, *Cortinarius*, *Lactarius*, *Russula*, and *Tricholoma* are symbionts of the forest trees.

He also turned his interest to the physiological interactions between the trees and their fungal partners. In pure culture experiments with mycorrhizal fungi, performed in liquid media after the addition of minute amounts of exudates from germinating seeds of pine and spruce, he found that these fungi required simple sugars for growth but were unable to decompose cellulose. He concluded that the mycorrhizal fungi obtain sugars from the tree roots, and that on the other hand, the fungi are especially important to the trees in transferring nitrogen from the soil to the roots. These earlier works and results were summarized by Melin in his monograph *Untersuchungen über die Bedeutung der Baummykorrhiza* (Jena 1925) which is a classic of mycological literature.

In 1930 Melin became professor of botany at the university of Uppsala, where he introduced the study of modern plant physiology. A large number of students,

not only from Sweden but also from abroad, were attracted by Melin and his ideas, and under his leadership an intense activity developed in the field of mycorrhiza and physiology of fungi. The laboratory facilities necessary for this development were created under his leadership in the modern institute for plant physiology at the university of Uppsala.

Melin quickly appreciated the possibilities of isotope tracer techniques in mycorrhizal research. In co-operation with H. Nilsson he used this method to prove the transfer of products of photosynthesis from pine seedlings to the fungal symbiont, as well as the transfer of nitrogen, phosphorus, and calcium from the fungus to the plant.

In his later years, Melin still retained his keen interest in the problems of the mycorrhiza, especially the stimulating effect of the plant roots on the growth and spore germination of mycorrhizal fungi. He concluded that this effect was due to exudates from the roots and made considerable efforts to elucidate the chemical nature of this so called 'M-factor'.

Besides his scientific activity, Melin was deeply interested and involved in the social life of the students both at his department and at his student union. His election in 1946 as inspector (master) of this union gave him the greatest pleasure and emphasises the affectionate esteem in which he was held by the students. Colleagues and pupils in several countries are deeply indebted to him for his outstanding leadership and his never-failing friendship and support.

G. Lindeberg

Scott R. Mazzur

SCOTT R. MAZZUR, well-known for her work on social and cultural factors affecting the transmission of hepatitis, died in Washington, DC on 2 March 1979.

Scott Mazzur was born on 30 September 1932 in Princeton, New Jersey and obtained her PhD in 1966 at the University of Pennsylvania working with Dr Werner Henle. Following this, she directed a viral diagnostic laboratory for the New Jersey State Department of Health until 1969 when she joined the laboratory of Dr Barry Blumberg at the Institute for Cancer Research in Philadelphia. In 1973, she was appointed Head of the Viral Epidemiology Section of American Red Cross Blood Services.

Scott Mazzur was fascinated by the interaction and mutual pressure exerted by evolution on human society and human disease. She recognized that a virus like rubella, which has no dormant state, must pass from person to person to maintain

itself and that this would be possible only in stable, larger communities which would be able to provide a sufficient number of new contacts. This means that many of the diseases that we know today have probably existed only for the last five or six thousand years when urban communities began to develop in Mesopotamia.

Much of her work was carried out in a series of field trips to a remote area of the Solomon Islands in the South Pacific where the incidence of hepatitis carriers is about five hundred times greater than in the United States. The unique feature of her approach was the prospective study of the total human population of the island, thereby eliminating the elements of chance and selection that have biased the statistical analysis of many studies. The simplicity of her approach was remarkable: she was entirely self-sufficient and did not take any assistants with her on her four field trips to the South Pacific between 1972 and 1977. She took a full census of the residents, drew blood samples and tested them in her field laboratory. She learned to speak Pidgin English and was perfectly at ease in communicating with 'my people' as she like to call them, even though head-hunting and cannibalism was practiced in these islands as recently as the second world war. She loved the islanders and her love was reciprocated by them, as shown by the fact that she was able to collect blood samples even from infants in the various communities.

Scott Mazzur used subtypes of hepatitis B surface antigen extensively as an epidemiological tool and had recently described a new antigenic marker 'e' and used it to subdivide other specificities. Her studies in the South Pacific provided new insights into hepatitis B virus epidemiology. The practice of separating male and female children in the Solomon Islands' culture enabled her to reject the hypothesis that there is a genetic basis for the HBsAg carrier state and to show that high frequency of carriers in males was due to a higher rate of infection, chiefly due to cultural practices. She confirmed earlier suggestions that children under six years have a tendency to become asymptomatic chronic carriers if infected with HBV, while older children and adults often react with an immune response.

Scott Mazzur saw her studies of hepatitis as being a model for other infections of similar epidemiology which might also mimic genetic patterns, for example, in the so-called polyclonal birth defects. This kind of distribution could result from the prolonged presence of infectious carriers combined with the limited dissemination of the infectious agent. Her work bridged anthropology and epidemiology and to each of these she brought new insights. She was a warm and generous friend who met with courage the adversities of her husband's death and her own illness.

G.A. Jamieson